



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 158135

TO: Jeanine Goldberg
Location: rem/2D15/2C70
Art Unit: 1634
Wednesday, July 13, 2005
Case Serial Number: 10/681199

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Goldberg,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523

This Page Blank (uspto)

STIC-Biotech/ChemLib

158136

mg

From: Goldberg, Jeanine
Sent: Tuesday, July 05, 2005 6:20 AM
To: STIC-Biotech/ChemLib
Subject: 10/681199- Search

Please search SEQ ID NO: 1 in all databases.
Please also do a mer search of SEQ ID NO: 1.

THANK YOU

Jeanine Anne Goldberg
1634
571-272-0743
REM 2D15

2070

RECEIVED
JUL -5 2005
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 8, 2005, 04:46:25 ; Search time 758 Seconds
(without alignments)
9863.632 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgctcttcaggtagcga.....gaacagaactaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	1263	9 ADB16964	ADB16964 Human DYX
2	1263	100.0	1993	9 ADB16965	ADB16965 cDNA sequ
3	1049	83.1	1641	10 ADC30210	ADC30210 Human nov
4	1002	79.3	1263	9 ADB16939	ADB16939 Pygmy chi
5	998	79.0	1559	11 ADM01890	ADM01890 Human cDN
6	960	76.0	1263	9 ADB16933	ADB16933 Chimpanze
7	859	68.0	1263	9 ADB16935	ADB16935 Gorilla D
8	696	55.1	1263	9 ADB16937	ADB16937 Orangutan
9	524	41.5	608	10 ADC32116	ADC32116 Human nov
10	326	25.8	488	9 ACH35463	ACH35463 Human end
11	270	21.4	458	9 ACH23091	ACH23091 Human adu
12	197	15.6	1383	5 AAS70018	AAS70018 DNA encod
13	184	14.6	49806	9 ADB16927	ADB16927 Human DYX
14	155	12.3	49939	9 ADB16928	ADB16928 Human DYX
15	126	10.0	164	3 AAC30498	AAC30498 Human sec
16	109	8.6	313	3 AAC26799	AAC26799 Human sec
17	60	4.8	174	3 AAA45298	AAA45298 Human sec
18	56	4.4	174	3 AAA45298	AAA45298 Human sec
19	33	2.6	1697	9 ADB16924	ADB16924 cDNA sequ
20	25	2.0	25	9 ADB16932	ADB16932 Reverse R

c	21	25	2.0	25	9	ADB16931	ADB16931 Forward R
	22	22	1.7	1316	2	AAC07163	AAC07163 Human lun
	23	22	1.7	1316	3	AAC79075	AAC79075 Human lun
	24	22	1.7	1316	4	AAD23150	AAD23150 Human lun
	25	22	1.7	1316	10	ADD67087	ADD67087 Human lun
	26	22	1.7	1316	10	ADDE87592	ADDE87592 Human lun
	27	22	1.7	1431	4	ABL01953	ABL01953 Drosophil
	28	22	1.7	5746	4	ABL01952	ABL01952 Drosophil
c	29	21	1.7	579	4	AAB09057	AAB09057 Human cDN
	30	21	1.7	2310	12	ADI81619	ADI81619 C. elegan
	31	21	1.7	310268	13	ABD32548	ABD32548 Human can
	32	20	1.6	480	4	AAI86685	AAI86685 Human pol
	33	20	1.6	481	6	ABL83144	ABL83144 Human ova
c	34	20	1.6	1339	10	ADF82325	ADF82325 Leukaemia
	35	20	1.6	2967	13	ADR08160	ADR08160 Full leng
	36	20	1.6	3075	13	ADR08053	ADR08053 Full leng
	37	20	1.6	3143	13	ADR66793	ADR66793 Human pro
	38	20	1.6	3143	13	ADR65890	ADR65890 Human pro
	39	20	1.6	3183	13	ADR08303	ADR08303 Full leng
	40	20	1.6	8073	6	ABL32754	ABL32754 Human imm
	41	20	1.6	9110	6	ABL34429	ABL34429 Human imm
	42	20	1.6	48551	6	AAS20800	AAS20800 Clostridi
c	43	20	1.6	84248	6	ABQ99651	ABQ99651 Human MS4
	44	20	1.6	110000	12	ADQ97138_3	Continuation (4 of
	45	20	1.6	110469	12	ADQ97337	ADQ97337 Human can

ALIGNMENTS

RESULT 1
ADB16964
ID ADB16964 standard; cDNA; 1263 BP.
XX
AC ADB16964;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human DYXC1 cDNA with single nucleotide polymorphisms.
XX
KW Gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW Chromosome 15q21; reading disability; phonological processing;
KW rapid naming; verbal short-term memory; single nucleotide polymorphism;
KW SNP.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 1..1263
FT /tag= a
FT /product= "DYXC1 protein"
FT replace(4,t)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT replace(572,a)
FT /tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1249,t)
FT /tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1259,g)
FT /tag= e
FT /standard_name= "Single nucleotide polymorphism"
XX
WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.

XX PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX DR WPI; 2003-646482/61.
XX DR P-PSDB; ADB16923.
XX PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
XX PT e.g. reading, phonological processing, rapid naming or verbal short-term
XX PT memory, or for diagnosing dyslexia or assessing the predisposition to
XX PT dyslexia.
XX PS Claim 1; Page 47; 135pp; English.
XX SS This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide
CC represents the coding sequence of human DYXC1 cDNA of the invention.
XX SQ Sequence 1263 BP; 491 A; 214 C; 275 G; 283 T; 0 U; 0 Other;

Query Match 100.0%; Score 1263; DB 9; Length 1263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCTCTTCAGGTAGGATTACAGCTGGCAGCAGACGAGACTGGCGTCTTCTGTCT 60
Db 1 ATGCCTCTTCAGGTAGGATTACAGCTGGCAGCAGACGAGACTGGCGTCTTCTGTCT 60

Qy 61 CTGCCCCCTCAAGCGGTGCTGCAGACACGACGCTGTTCTGCACGGAACCTATCTG 120
Db 61 CTGCCCCCTCAAGCGGTGCTGCAGACACGACGCTGTTCTGCACGGAACCTATCTG 120

Qy 121 AAGGTCAACTTCTCCATTTTATTTTGAAGCAATTTCTTATGTCTCCATAGCAGTGA 180
Db 121 AAGGTCAACTTCTCCATTTTATTTTGAAGCAATTTCTTATGTCTCCATAGCAGTGA 180

Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCACTTTGTATATAAAGAGCG 240
Db 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCACTTTGTATATAAAGAGCG 240

Qy 241 GCCATGTGGGAGACCTTTCTGTGACGGGTGTCACAAAGAGATGATGCAAGAAATAGA 300
Db 241 GCCATGTGGGAGACCTTTCTGTGACGGGTGTCACAAAGAGATGATGCAAGAAATAGA 300

Qy 301 GAAAAATCTATTTTACAAGCAACAAGAGAGACCAAGAGCTACAGAGCAAAAGCTGCA 360
Db 301 GAAAAATCTATTTTACAAGCAACAAGAGAGACCAAGAGCTACAGAGCAAAAGCTGCA 360

Qy 361 GCAAAGCGGGAAGATCAAAAATAGCACAATAGTGTCTATGATGAAGTTGAAGAAGAG 420
Db 361 GCAAAGCGGGAAGATCAAAAATAGCACAATAGTGTCTATGATGAAGTTGAAGAAGAG 420

Qy 421 AGGAAAAATATAGAGATATCAAAAGAAATCAACGGAATCAAGCACTTAAAGCAATTGAA 480
Db 421 AGGAAAAATATAGAGATATCAAAAGAAATCAACGGAATCAAGCACTTAAAGCAATTGAA 480

Qy 481 GCCTGGAAGAAATATCAAAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 540
Db 481 GCCTGGAAGAAATATCAAAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 540

Qy 541 TTATGTCAAAAGAAAGCAAAATTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 TTATGTCAAAAGAAAGCAAAATTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTCTGCCAAAGGGAGAAATTCAGAGAAATATATTT 660

Db 601 ACTAGAAATTTGGCATCTAGAAATCTTCTGCCAAAGGGAGAAATTCAGAGAAATATATTT 660

Qy 661 ACTGAGAAAGTTAAAGGAAGACAGTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720

Db 661 ACTGAGAAAGTTAAAGGAAGACAGTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720

Qy 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCACAAGTAGCAGAGAGAG 780

Db 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCACAAGTAGCAGAGAGAG 780

Qy 781 GAGTGGCTACACAAACAAAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAATCTT 840

Db 781 GAGTGGCTACACAAACAAAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAATCTT 840

Qy 841 TGGATTTAAAG 900

Db 841 TGGATTTAAAG 900

Qy 901 TTTGCAACGGAACAACTATTTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960

Db 901 TTTGCAACGGAACAACTATTTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960

Qy 961 AATAAGATGCGCACTATTTGTAATTTGAACCGGGCTGCTTGCACCTTAAACAACTTA 1020

Db 961 AATAAGATGCGCACTATTTGTAATTTGAACCGGGCTGCTTGCACCTTAAACAACTTA 1020

Qy 1021 CACAGGCTATTTGAAGATTTCTTCTAGGCACTGGAATTTATGATGCCACCTGTTACAGAC 1080

Db 1021 CACAGGCTATTTGAAGATTTCTTCTAGGCACTGGAATTTATGATGCCACCTGTTACAGAC 1080

Qy 1081 AATGCTAATCAAGAAATGAAGGACACATGTACAGCTGGAACAGCATTTCTGCAACTAGAA 1140

Db 1081 AATGCTAATCAAGAAATGAAGGACACATGTACAGCTGGAACAGCATTTCTGCAACTAGAA 1140

Qy 1141 TTGTATGTAGAACGCTTACAGGATTTATGAACGCGCACTTAAAGATTTGATCCATCCAA 1200

Db 1141 TTGTATGTAGAACGCTTACAGGATTTATGAACGCGCACTTAAAGATTTGATCCATCCAA 1200

Qy 1201 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAATTCAGGAAACAGAACTTAAATCT 1260

Db 1201 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAATTCAGGAAACAGAACTTAAATCT 1260

Qy 1261 TAA 1263

Db 1261 TAA 1263

RESULT 2
ADB16965
ID ADB16965 standard; cDNA; 1993 BP.
XX ADB16965;
AC AC
XX XX
DT 20-NOV-2003 (first entry)
XX cDNA sequence of the human DYXC1 mRNA.
XX gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW chromosome 15q21; reading disability; phonological processing;
KW rapid naming; verbal short-term memory; SNP;
KW single nucleotide polymorphism.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH variation replace(205,t)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 369..11631
FT /*tag= b
FT /product= "DYXC1 protein"
XX XX

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 15q21.3; gene; ss.
OS Homo sapiens.
XX WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Dmanac RT;
XX WPI; 2003-371981/35.
DR P-PSDB; ADC31181.
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX Claim 1; SEQ ID NO 292; 1185pp; English.
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1641 BP; 614 A; 272 C; 331 G; 424 T; 0 U; 0 Other;
Query Match 83.1%; Score 1049; DB 10; Length 1641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACGACGACGCTTCTGTCT 60
DB 104 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACGACGACGCTTCTGTCT 163
QY 61 CTGCCCTCAAAGGCGTGTGCGTCAGACACGCGGCGTGTCTGCACGAAAACTATCTG 120

DB 164 CTGCCCTCAAAGGCGTGTGCGTCAGACACGCGGCGTGTCTGCACGAAAACTATCTG 223
QY 121 AAGGTCAACTTCTCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
DB 224 AAGGTCAACTTCTCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 283
QY 181 AGCAGCAAAAGCAAAGATTGGGAATGACACCATTTGTCTTCACTTTGTATATAAAAAAGAGCG 240
DB 284 AGCAGCAAAAGCAAAGATTGGGAATGACACCATTTGTCTTCACTTTGTATATAAAAAAGAGCG 343
QY 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAAGAAATPAGA 300
DB 344 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAAGAAATPAGA 403
QY 301 GAAAAATCTATTTTACAGCACAAAGAGAGAGCAAAAGAGAGCTACAGNAGCAAAAGCTGCA 360
DB 404 GAAAAATCTATTTTACAAAGCACAAAGAGAGAGCAAAAGAGAGCTACAGNAGCAAAAGCTGCA 463
QY 361 GCAAGCGGGAAGATCAAAATACGCACCTAAGTGTCTATGATGAAGATTGAAAGAAAGAGAG 420
DB 464 GCAAGCGGGAAGATCAAAATACGCACCTAAGTGTCTATGATGAAGATTGAAAGAAAGAGAG 523
QY 421 AGGAAAAAATAGAGATATGAAAGAAAAATGAAACGGATATAAAAGCCACTAAAGCATTTGAA 480
DB 524 AGGAAAAAATAGAGATATGAAAGAAAAATGAAACGGATATAAAAGCCACTAAAGCATTTGAA 583
QY 481 GCCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAGAAA 540
DB 584 GCCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAGAAA 643
QY 600 TTATGTCAAAAAGAAAGCAAAATTTAAAGAAAGGAGAAAAAATAAATAATTAAGAGTCTT 600
DB 703 TTATGTCAAAAAGAAAGCAAAATTTAAAGAAAGGAGAAAAAATAAATAATTAAGAGTCTT 703
QY 660 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 660
DB 763 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 763
QY 720 ACTGAGAAAGTTAAAGGAAGACAGTATTCCTGCTCTCTGCTCTGTTGGCAGTATTAATATC 720
DB 823 ACTGAGAAAGTTAAAGGAAGACAGTATTCCTGCTCTCTGCTCTGTTGGCAGTATTAATATC 823
QY 780 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTAATCAACAAGTAGCAGAGAGAG 780
DB 883 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTAATCAACAAGTAGCAGAGAGAG 883
QY 840 GAGTGGCTACACAAACAGCTGAGGCAAGAGCAATGAATACTGACATAGCTGAACCTT 840
DB 943 GAGTGGCTACACAAACAGCTGAGGCAAGAGCAATGAATACTGACATAGCTGAACCTT 943
QY 900 TGGGATTTAAAG 900
DB 1003 TGGGATTTAAAG 1003
QY 960 TTTGCAAGGAG 960
DB 1063 TTTGCAAGGAG 1063
QY 961 AATAAGATGACCATTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1123 AATAAGATGACCATTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123
QY 1021 CACAAGGCTATTGAGAGATTTCTTCTAAGGC 1049
DB 1124 CACAAGGCTATTGAGAGATTTCTTCTAAGGC 1152
RESULT 4
ADB16939
ID ADB16939 standard; cDNA; 1263 BP.
XX
AC ADB16939;

XX ADM01890;
AC 20-MAY-2004 (first entry)
DT Human cDNA of the invention SEQ ID NO:575.
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX Homo sapiens.
XX EPI347046-A1.
XX 24-SEP-2003.
XX 12-APR-2002; 2002EP-00008400.
XX 22-MAR-2002; 2002JP-00137785.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Makamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-723558/69.
DR P-PSDB; ADM04333.
DR
XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 575; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.
XX
SQ Sequence 1559 BP; 592 A; 252 C; 309 G; 406 T; 0 U; 0 Other;
Query Match 79.0%; Score 998; DB 11; Length 1559;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACGAGACGTGCGGTCTTTCTGTCT 60
DB 43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACGAGACGTGCGGTCTTTCTGTCT 102
QY 61 CTGCCCTCAAAAGGGTGTGGTTCAGACACGACGCTGTTCTGCAGGAAAACATCTGTG 120
DB 103 CTGCCCTCAAAAGGGTGTGGTTCAGACACGACGCTGTTCTGCAGGAAAACATCTGTG 162
QY 121 AAGGTCAACTTTCCTCCATTTTATTTTCAGGCACTTTCTTATGCTCCCATAGACGATGAG 180
DB 163 AAGGTCAACTTTCCTCCATTTTATTTTCAGGCACTTTCTTATGCTCCCATAGACGATGAG 222
QY 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTTCACCTTGTATATAAAAAAGAGCG 240
DB 223 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTTCACCTTGTATATAAAAAAGAGCG 282
QY 241 GCCATGTGGGAGACCTTTCTGTGCGGGTGTTCAGCAAGATGATGATCAAGAAATTAGA 300
DB 283 GCCATGTGGGAGACCTTTCTGTGCGGGTGTTCAGCAAGATGATGATCAAGAAATTAGA 342
QY 301 GAAAAATCTATTTTACAAGCAACAGAGAGAGCAAAAGAGCTTACAGAGCAAGAAAGCTGCA 360
|||||

DB 343 GAAAAATCTATTTTACAAGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 402
QY 361 GCAAAGCGGGAAGATCAAAATATAGCACTAAGTGTCTCATGTAGATTGAGAGAGAGAG 420
DB 403 GCAAAGCGGGAAGATCAAAATATAGCACTAAGTGTCTCATGTAGATTGAGAGAGAGAG 462
QY 421 AGCAAAAAAATAGAGATATGAAAGAAAAATGAAACGGATATAAAGCCACTTAAGCATTCGAA 480
DB 463 AGCAAAAAAATAGAGATATGAAAGAAAAATGAAACGGATATAAAGCCACTTAAGCATTCGAA 522
QY 481 GCCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 540
DB 523 GCCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 582
QY 541 TTATGTCAAAAAGAAAGCAAAATTAAGAGAGGAGAAATAAATAAATAAAGAGTCTT 600
DB 583 TTATGTCAAAAAGAAAGCAAAATTAAGAGAGGAGAAATAAATAAATAAAGAGTCTT 642
QY 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGAGAGAAATTCAGAAAAATATATTT 660
DB 643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGAGAGAAATTCAGAAAAATATATTT 702
QY 661 ACTGAGAGAGTTAAAGGAGAGACAGTATTCCTGCTCCTGCTCTGTTGGCAGTATTAATAATC 720
DB 703 ACTGAGAGAGTTAAAGGAGAGACAGTATTCCTGCTCCTGCTCTGTTGGCAGTATTAATAATC 762
QY 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGAGAG 780
DB 763 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGAGAG 822
QY 781 GAGTGGCTACACAAAACAAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAACCTT 840
DB 823 GAGTGGCTACACAAAACAAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAACCTT 882
QY 841 TGGCATTTAAAGAGAGAGAGAAAGAACCCAGATGTTGAGAGATATAAGGAGAGAGAAATTTG 900
DB 883 TGGCATTTAAAGAGAGAGAGAGAAAGAACCCAGATGTTGAGAGATATAAGGAGAGAGAAATTTG 942
QY 901 TTTGCAACGGAAGAACTATTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960
DB 943 TTTGCAACGGAAGAACTATTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 1002
QY 961 AATAAGATGCACTATTTGATTTTGAACCGGCTGCTTGGCCACCTTAAACCTAAAAAACTTA 1020
DB 1003 AATAAGATGCACTATTTGATTTTGAACCGGCTGCTTGGCCACCTTAAACCTAAAAAACTTA 1062
QY 1021 CACAAGGCTATTTGAAGATTTCTTTAAGGC 1049
DB 1063 CACAAGGCTATTTGAAGATTTCTTTAAGGC 1091
RESULT 6
ADBI6933
ID ADBI6933 standard; cDNA; 1263 BP.
XX
AC ADBI6933;
XX 20-NOV-2003 (first entry)
DT
XX Chimpanzee DYX1C1 cDNA sequence.
DE
XX
KW gene; ss; chimpanzee; DYX1C1; dyslexia; neurological disorder;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Pan troglodytes.
XX
XX Key Location/Qualifiers
FT CDS 1..1263
FT /*tag= a
FT /product= "DYX1C1 protein"
XX
PN WO2003068814-A1.

XX WO2003068814-A1.
 PN 21-AUG-2003.
 PD 12-FEB-2003; 2003WO-FI000110.
 XX 12-FEB-2003; 2002US-0355782P.
 PR 12-FEB-2002; 2002US-0355782P.
 XX (LION) LICENTIA LTD.
 XX Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
 PI WPI: 2003-646482/61.
 XX P-PSDB; ADB16936.
 DR New isolated, purified DYX1C1 nucleic acid for studying brain processes,
 PT e.g. reading, phonological processing, rapid naming or verbal short-term
 PT memory, or for diagnosing dyslexia or assessing the predisposition to
 PT dyslexia.
 XX Claim 29; Page 119-121; 135pp; English.
 PS This invention relates to a novel isolated human gene DYX1C1 that is
 CC functionally related to dyslexia, more particularly it describes single
 CC nucleotide polymorphisms thought to predispose an individual in to
 CC developing dyslexia. This is a neurological disorder with a genetic basis
 CC (DYX1C1 has been isolated to chromosome 15q21), which manifests itself as
 CC a specific reading disability. Specifically, DYX1C1 is can be useful in
 CC study of brain processes such as reading, phonological processing, rapid
 CC naming and verbal short-term memory. Accordingly, the present invention
 CC describes methods and materials for analysing allelic variations in the
 CC DYX1C1 gene, and also provides DYX1C1 as an antigen for the production of
 CC antibodies used in the diagnosis of dyslexia. This polynucleotide
 CC sequence is the gorilla DYX1C1 cDNA homologous to the human DYX1C1 gene of
 CC the invention.
 XX Sequence 1263 BP; 491 A; 212 C; 275 G; 285 T; 0 U; 0 Other;
 SQ

Query Match 68.0%; Score 859; DB 9; Length 1263;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1209; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 48 GGTCTTTCTGCTCTGCCCTCAAGGCGTGTGCTGAGACACGACGCGTGTCTGCAC 107
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 108 GGAAGAACTATCTGAAGGTCAACTTTCTCTCCATTTTATTTGAGGCATTTCTTTATGCTCC 167
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 168 CATAGACGATGAGAGCAGCAAGCAAGATGGGAATGACACCATTTGCTTCACTGTGA 227
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 228 TAAAGAGAGCGGCCATGTGGGAGACCTTTCTGTGACGGGTGTGACAAAGAGATGAT 287
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 288 GCAAGAAATAGAGAAAATCTATTTTACAGCACAAGAGAGAGCAAAAGAGCTACAGA 347
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 348 AGCAAAAGCTGCAGCAAGCGGGAGATCAAAAATACGCATTAAGTTCATGATGAAGAT 407
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 408 TGAAGAGACAGAGGAGAAAATAGAGATATGAAGAAAATGAACGGATAAAGCCAC 467
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 468 TAAAGCAATTTGGAAGCTGGAAGAAATATCAAGAAAAGCTGAGGAGCAAAAATTTCA 527
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 468 TAAAGCAATTTGGAAGCTGGAAGAAATATCAAGAAAAGCTGAGGAGCAAAAATTTCA 527
 QY 528 GAGAGAGAGAAAATTTATGTCAAAAAGAAAAGCAAAATTAAGAGAGAAAAGAAAATATAA 587
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 528 GAGAGAGAGAGATTTATGTCAAAAAGAAAAGCAAAATTAAGAGAGAAAAGAAAATATAA 587
 QY 588 ATATAAGAGTCTTACTAGAAAATTTGGCATCTAGAAAATCTTGTCTCCAAAAGGAGAAATTC 647
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 648 AGAAAATATATTTACTGAGAAATTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
 DB 648 AGAAAATATATTTACTGAGAAATTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
 QY 708 CAGTATTTAAATCAACTTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGCTGAATCACAAGT 767
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 708 CAGTATTTAAATCAACTTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGCTGAATCACAAGT 767
 QY 768 AGCAGAGAGAGAGAGTGGCTACACAAACAAGCTGAGGACGAGAGCAATGAATCTGA 827
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 768 AGCAGAGAGAGAGAGTGGCTACACAAACAAGCTGAGGACGAGAGCAATGAATCTGA 827
 QY 828 CATAGCTGAACCTTTGCGATTTAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 887
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 828 CATAGCTGAACCTTTGCGATTTAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 887
 QY 888 AGGAAACAAATTTGTTGCAACGAGAAACTATTTGGCAGCTATCAATGCATATATTTAGC 947
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 888 AGGAAACAAATTTGTTGCAACGAGAAACTATTTGGCAGCTATCAATGCATATATTTAGC 947
 QY 948 CATAGAAGTAAATTAAGATGCCACTTATTTGTAATTTGAACCGGCTCTTGGCCACTTAA 1007
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 948 CATAGAAGTAAATTAAGATGCCACTTATTTGTAATTTGAACCGGCTCTTGGCCACTTAA 1007
 QY 1008 ACTAAAAAATTTACACAAGGCTATTTGAAGATTTCTTCTAAGGCACTGGAAATTTATGATGCC 1067
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1008 ACTAAAAAATTTACACAAGGCTATTTGAAGATTTCTTCTAAGGCACTGGAAATTTATGATGCC 1067
 QY 1068 ACCTGTTTACAGACAACTTAATGCAAGATGAAGGACATGTAGACGTGGAACAGCATTT 1127
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1068 ACCTGTTTACAGACAACTTAATGCAAGATGAAGGACATGTAGACGTGGAACAGCATTT 1127
 QY 1128 CTGTCAACTAGAAATTTGATGTAGAAAGCCTACAGGATTTATGAAGCGGCACCTTAAGATTGA 1187
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1128 CTGTCAACTAGAAATTTGATGTAGAAAGCCTACAGGATTTATGAAGCGGCACCTTAAGATTGA 1187
 QY 1188 TCCATCCAAACAAATTTGATCAAAATTTGATGTAGAGAGATTTGGAATGTAATTTCAAGGAC 1247
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1188 TCCATCCAAACAAATTTGATCAAAATTTGATGTAGAGAGATTTGGAATGTAATTTCAAGGAC 1247
 QY 1248 AGAAGTAAATCTTAA 1263
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1248 AGAAGTAAATCTTAA 1263

RESULT 8
 ADB16937
 ID ADB16937 standard; cDNA; 1263 BP.
 XX
 AC ADB16937;
 AC AC
 XX AC
 DT 20-NOV-2003 (first entry)
 XX
 DE Orangutan DYX1C1 cDNA sequence.
 XX
 KW gene; ss; orangutan; DYX1C1; dyslexia; neurological disorder;
 KW reading disability; phonological processing; rapid naming;
 KW verbal short-term memory.
 XX
 OS Pongo pygmaeus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1263

[illegible]

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 15q21.3; ss.
XX Homo sapiens.
XX WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
DR P-PSDB; ADC32883.
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX Example 2; SEQ ID NO 2198; 1185pp; English.
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig
CC sequence used in an example of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 608 BP; 230 A; 103 C; 135 G; 140 T; 0 U; 0 Other;
SQ Query Match 41.5%; Score 524; DB 10; Length 608;
Best Local Similarity 100.0%; Pred. No. 1e-239;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 524 TTCAGAGAGAGAGAAATATGTCAAAAGAAAAGCAAAATTAAGAGAGAGAGAAAAA 583
DB 1 TTCAGAGAGAGAGAAATATGTCAAAAGAAAAGCAAAATTAAGAGAGAGAGAAAAA 60
QY 584 TAAATATAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGTCCAAAAGGAGAA 643

Db 61 TAAATATAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGTCCAAAAGGAGAA 120
QY 644 ATTCAGAAAATATATTTACTGAGAAAGTAAAGGAGAGACAGTATTTCTGCTCTCGCTCTG 703
Db 121 ATTCAGAAAATATATTTACTGAGAAAGTAAAGGAGAGACAGTATTTCTGCTCTCGCTCTG 180
QY 704 TTGGCAGTATTAATAATCAACTTTTACCCCTCGAGTATTTCCAAACAGCTCTTGTGTAATCAC 763
Db 181 TTGGCAGTATTAATAATCAACTTTTACCCCTCGAGTATTTCCAAACAGCTCTTGTGTAATCAC 240
QY 764 AAGTAGCAGAGAGAGGAGAGTGGCTACACAAAACAAGCTGAGGCCAGAGAGCAATGAATA 823
Db 241 AAGTAGCAGAGAGAGGAGAGTGGCTACACAAAACAAGCTGAGGCCAGAGAGCAATGAATA 300
QY 824 CTGACATAGCTGAACCTTTTGGGATTTTAAAGAGAGAGAAAAGAACCCAGAGATGGTTGAAG 883
Db 301 CTGACATAGCTGAACCTTTTGGGATTTTAAAGAGAGAGAAAAGAACCCAGAGATGGTTGAAG 360
QY 884 ATAAGGAGAGAGAGAGTGGCTTTCACACGGAACCACTATTTGGCAGCTATCAATGSCATATAATT 943
Db 361 ATAAGGAGAGAGAGAGTGGCTTTCACACGGAACCACTATTTGGCAGCTATCAATGSCATATAATT 420
QY 944 TAGCCATAAGACTAAATAAAGATGCCACTATTTGTATTTGAACCGGGCTCTTGGCCACC 1003
Db 421 TAGCCATAAGACTAAATAAAGATGCCACTATTTGTATTTGAACCGGGCTCTTGGCCACC 480
QY 1004 TAAAACTAAAAAACTTACACAGGCTATTTGAAGATTTCTTCTAAG 1047
Db 481 TAAAACTAAAAAACTTACACAGGCTATTTGAAGATTTCTTCTAAG 524
RESULT 10
ACH35463
ID ACH35463 standard; cDNA; 488 BP.
AC ACH35463;
XX 13-OCT-2003 (first entry)
DT Human endothelial cell cDNA #3596.
XX Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
OS US2003073623-A1.
PN 17-APR-2003.
PD 30-JUL-2001; 2001US-00918995.
PF 30-JUL-2001; 2001US-00918995.
PR 30-JUL-2001; 2001US-00918995.
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
PI WPI; 2003-615964/58.
DR New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX Claim 1; SEQ ID NO 22675; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX

SQ Sequence 488 BP; 141 A; 105 C; 126 G; 116 T; 0 U; 0 Other;

Query Match 25.8%; Score 326; DB 9; Length 488;

Best Local Similarity 100.0%; Pred. No. 4.1e-145;

Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GGGTCAGAGACACGGACGGTTCGACGGAACCTATCTGAAGTCAACTTCTCTCCAT 139
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 140 TTTTATTTGAGGCATTCTTTATGCTCCATAGACGATGAGCAGCAAGCAAGATTG 199
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 133 TTTTATTTGAGGCATTCTTTATGCTCCATAGACGATGAGCAGCAAGCAAGATTG 192
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 200 GGAATGACACCATTTCTTCACTTGTATATAAAGAGCGGCCATGTGGAGACCTTT 259
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 193 GGAATGACACCATTTCTTCACTTGTATATAAAGAGCGGCCATGTGGAGACCTTT 252
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 260 CTGTACGGGTGTTGACAAAGAGATGATGCAAGAACTATAGAAAATCTTTTACAG 319
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 253 CTGTACGGGTGTTGACAAAGAGATGATGCAAGAACTATAGAAAATCTTTTACAG 312
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 320 CACAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCAGCAAAAGCGGGAAGATCAA 379
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 313 CACAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCAGCAAAAGCGGGAAGATCAA 372
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 380 AATACGCACTAAGTGTCTATGATGAAG 405
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 373 AATACGCACTAAGTGTCTATGATGAAG 398.

RESULT 11

ACH23091 ID ACH23091 standard; cDNA; 458 BP.

XX AC ACH23091;

XX DT 13-OCT-2003 (first entry)

XX Human adult ovary cDNA #1471.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX PD 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRNA/) DRMANAC R T.

XX (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

PS Claim 1; SEQ ID NO 10303; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX

SQ Sequence 458 BP; 170 A; 86 C; 97 G; 102 T; 0 U; 3 Other;

Query Match 21.4%; Score 270; DB 9; Length 458;

Best Local Similarity 100.0%; Pred. No. 2.3e-118;

Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 773 AAGAGGAGGAGTGGCTACACAAAGCTGAGGCGACGAAGAGCAATGATCTGACATAG 832
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 189 AAGAGGAGGAGTGGCTACACAAAGCTGAGGCGACGAAGAGCAATGATCTGACATAG 248
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 833 CTGAACCTTTGCGATTTAAAGAGAGAAAGAACCCAGATGGTTGAAGGATAAGGAA 892
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 249 CTGAACCTTTGCGATTTAAAGAGAGAAAGAACCCAGATGGTTGAAGGATAAGGAA 308
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 893 ACAAAATTTGTCACACGGAACCTATTTGGCAGCTATCAATGCATATATTTAGCCATAA 952
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 309 ACAAAATTTGTCACACGGAACCTATTTGGCAGCTATCAATGCATATATTTAGCCATAA 368
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 953 GACTAAATATAAGATGCGCACTATTGTTATTTGAACCGGGCTGCTGCCAACCCTAAACTAA 1012
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 369 GACTAAATATAAGATGCGCACTATTGTTATTTGAACCGGGCTGCTGCCAACCCTAAACTAA 428
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1013 AAAAATTACACAGGCTATTGAAGATCTT 1042
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 429 AAAAATTACACAGGCTATTGAAGATCTT 458

RESULT 12

AAAS70018 ID.

XX ID. AAAS70018 standard; cDNA; 1383 BP.

XX AC AAAS70018;

XX DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #5822.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Dmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG05831.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 1; SEQ ID NO 5822; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1383 BP; 399 A; 343 C; 337 G; 304 T; 0 U; 0 Other;
 SQ
 Query Match 15.6%; Score 197; DB 5; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 1.8e-83;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 172 GACGATGAGCAGCAAGCAAGATTGGGATGACACCATTTCTTCCACCTTGTATAAA 231
 DB 709 GACGATGAGCAGCAGCAAGCAAGATTGGGATGACACCATTTCTTCCACCTTGTATAAA 768
 OY 232 AAAGAAGCGGCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGATGATGCAA 291
 DB 769 AAAGAAGCGGCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGATGATGCAA 828
 OY 292 AGAATTAGAGAAATCTATTTCACAGCAAGAGAGAGCAAGAGAGCTTACAGAGCA 351
 DB 829 AGAATTAGAGAAATCTATTTCACAGCAAGAGAGAGCAAGAGAGCTTACAGAGCA 888
 OY 352 AAAGCTGCAGCAAGCG 368
 DB 889 AAAGCTGCAGCAAGCG 905
 RESULT 13
 ADB16927
 ID ADB16927 standard; DNA; 49806 BP.

XX ADB16927;
 XX 20-NOV-2003 (first entry)
 DE Human DYXC1 DNA, chromosomal gene region nucleotides 50001-100000.
 XX ds; human; DYXC1; dyslexia; neurological disorder; chromosome 15q21;
 KW reading disability; phonological processing; rapid naming;
 KW verbal short-term memory.
 XX Homo sapiens.
 XX WO2003068814-A1.
 XX 21-AUG-2003.
 XX 12-FEB-2003; 2003WO-FI000110.
 PR 12-FEB-2002; 2002US-0355782P.
 XX (LICN) LICENTIA LTD.
 XX Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
 XX WPI; 2003-646482/61.
 XX New isolated, purified DYXC1 nucleic acid for studying brain processes,
 PT e.g. reading, phonological processing, rapid naming or verbal short-term
 PT memory, or for diagnosing dyslexia or assessing the predisposition to
 PT dyslexia.
 XX Claim 27; Page 69-83; 135pp; English.
 PS This invention relates to a novel isolated human gene DYXC1 that is
 CC functionally related to dyslexia, more particularly it describes single
 CC nucleotide polymorphisms thought to predispose an individual in to
 CC developing dyslexia. This is a neurological disorder with a genetic basis
 CC (DYXC1) has been isolated to chromosome 15q21, which manifests itself as
 CC a specific reading disability. Specifically, DYXC1 is can be useful in
 CC study of brain processes such as reading, phonological processing, rapid
 CC naming and verbal short-term memory. Accordingly, the present invention
 CC describes methods and materials for analysing allelic variations in the
 CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
 CC antibodies used in the diagnosis of dyslexia. This polynucleotide is the
 CC partial genomic sequence of the human DYXC1 chromosomal region
 CC (nucleotides 50001-100000) of the invention.
 XX Sequence 49806 BP; 12975 A; 10105 C; 10577 G; 15977 T; 0 U; 172 Other;
 SQ
 Query Match 14.6%; Score 184; DB 9; Length 49806;
 Best Local Similarity 99.6%; Pred. No. 2.8e-77;
 Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 404 AGATTGAAGAAGAGAGAGAGGAAAGAAATATGAAGATATGAAGAAATGAACGATATAAG 463
 DB 32444 AGATTGAAGAAGAGAGAGGAGGAAAGAAATATGAAGATATGAAGAAATGAACGATATAAG 32503
 OY 464 CCACATAAGCATTCGGAAGCCCTGGAAAGAAATATCAAGAAAGAGCTGAGAGCAAAAAAAA 523
 DB 32504 CCACATAAGCATTCGGAAGCCCTGGAAAGAAATATCAAGAAAGAGCTGAGAGCAAAAAAAA 32563
 OY 524 TTTCAGAGAGAGAGAAATATGTCAAAAGAGAGCAAAATTAAGAAAGAGAGAGAGAGAGAGAG 583
 DB 32564 TTTCAGAGAGAGAGAAATATGTCAAAAGAGAGCAAAATTAAGAAAGAGAGAGAGAGAGAG 32623
 OY 584 TAAATATAGAGCTTCTACTAGAAATTTGGCATCTAGAAATCTTGTCTCCAAAAG 638
 DB 32624 TAAATATAGAGCTTCTACTAGAAATTTGGCATCTAGAAATCTTGTCTCCAAAAG 32678
 RESULT 14
 ADB16928

ID	ADBI6928 standard; DNA; 49939 BP.
XX	
AC	ADBI6928;
XX	
DT	20-NOV-2003 (first entry)
XX	
DB	Human DYX1 DNA, chromosomal gene region nucleotides 100001-150000.
XX	
KW	ds; human; DYX1; dyslexia; neurological disorder; chromosome 15q21;
KW	reading disability; phonological processing; rapid naming;
KW	verbal short-term memory.
XX	
OS	Homo sapiens.
XX	
XX	W02003068814-A1.
PN	
XX	
PD	21-AUG-2003.
XX	
PF	12-FEB-2003; 2003WO-FI000110.
XX	
PR	12-FEB-2002; 2002US-0355782P.
XX	
PA	(LICN) LICENTIA LTD.
XX	
PI	Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX	
DR	WPI; 2003-646482/61.
XX	
PT	New isolated, purified DYX1 nucleic acid for studying brain processes,
PT	e.g. reading, phonological processing, rapid naming or verbal short-term
PT	memory, or for diagnosing dyslexia or assessing the predisposition to
PT	dyslexia.
XX	
PS	Claim 27; Page 83-97; 135pp; English.
XX	
CC	This invention relates to a novel isolated human gene DYX1 that is
CC	functionally related to dyslexia, more particularly it describes single
CC	nucleotide polymorphisms thought to predispose an individual in to
CC	developing dyslexia. This is a neurological disorder with a genetic basis
CC	(DYX1 has been isolated to chromosome 15q21), which manifests itself as
CC	a specific reading disability. Specifically, DYX1 is can be useful in
CC	a study of brain processes such as reading, phonological processing, rapid
CC	naming and verbal short-term memory. Accordingly, the present invention
CC	describes methods and materials for analyzing allelic variations in the
CC	DYX1 gene, and also provides DYX1 as an antigen for the production of
CC	antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC	partial genomic sequence of the human DYX1 chromosomal region
CC	(nucleotides 100001-150000) of the invention.
XX	
SQ	Sequence 49939 BP; 13786 A; 10910 C; 11127 G; 14116 T; 0 U; 0 Other;
	Query Match 12.3%; Score 155; DB 9; Length 49939;
	Best Local Similarity 100.0%; Pred. No. 2e-63;
	Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	894 CAAATTGTTTGGCAACGGAAAACCTATTGGCAGCTATCATGATATATTAGCCATAAG 953
Db	14682 CAAATTGTTTGGCAACGGAAAACCTATTGGCAGCTATCATGATATATTAGCCATAAG 14741
Qy	954 ACTAATAATAAGATGCCACTATTGTATTGTAACCGGGCTGTTGCCACTTAAACTAAA 1013
Db	14742 ACTAATAATAAGATGCCACTATTGTATTGTAACCGGGCTGTTGCCACTTAAACTAAA 14801
Qy	1014 AAACCTTACACAGGCTATTGAAGATTTCTTCTAAGG 1048
Db	14802 AAACCTTACACAGGCTATTGAAGATTTCTTCTAAGG 14836

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 04:50:29 ; Search time 5660 Seconds
(without alignments)
10812.533 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggtagcga.....gaacagaactaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	1993	9 AF337549	Homo sapi
2	998	79.0	1559	6 AX833451	AX833451 Sequence
3	998	79.0	1559	9 AK095201	AK095201 Homo sapi
4	997	78.9	1468	9 BC062564	BC062564 Homo sapi
5	407	32.2	489	6 AR413597	AR413597 Sequence
6	407	32.2	489	6 AX970431	AX970431 Sequence
7	407	32.2	489	6 BD109150	BD109150 EST and e
8	235	18.6	114149	9 AC022083	AC022083 Homo sapi
9	214	16.9	464	6 CQ685075	CQ685075 Sequence
10	184	14.6	43886	2 AC021668	AC021668 Homo sapi
11	184	14.6	151133	9 AC013355	AC013355 Homo sapi
12	150	11.9	208181	2 AC016527	AC016527 Homo sapi
13	147	11.6	147	9 AY178583S2	AY178583S2 Pan trogl
14	147	11.6	147	9 AY178583S5	AY178583S5 Pan trogl
15	147	11.6	147	9 AY178592S2	AY178592S2 Pan panis
16	147	11.6	147	9 AY178592S5	AY178592S5 Pan panis
17	147	11.6	147	9 AY178601S2	AY178601S2 Gorilla g
18	147	11.6	147	9 AY178601S5	AY178601S5 Gorilla g
19	147	11.6	147	9 AY178610S2	AY178610S2 Pongo pyg

20	144	11.4	147	9	AY178610S5	AY178610S5 Pongo pyg
21	138	10.9	153	9	AY178583S7	AY178583S7 Pan trogl
22	138	10.9	153	9	AY178592S7	AY178592S7 Pan panis
23	138	10.9	153	9	AY178601S7	AY178601S7 Gorilla g
24	138	10.9	153	9	AY178610S7	AY178610S7 Pongo pyg
25	134	10.6	1187	6	C0732459	C0732459 Sequence
26	129	10.2	231	9	AY178583S4	AY178583S4 Pan trogl
27	129	10.2	231	9	AY178592S4	AY178592S4 Pan panis
28	126	10.0	164	6	AX918710	AX918710 Sequence
29	126	10.0	164	6	BD054243	BD054243 Sequence
30	117	9.3	123	9	AY178592S1	AY178592S1 Pan panis
31	117	9.3	123	9	AY178610S1	AY178610S1 Pongo pyg
32	113	8.9	135	9	AY178583S3	AY178583S3 Pan trogl
33	113	8.9	135	9	AY178592S3	AY178592S3 Pan panis
34	113	8.9	135	9	AY178601S3	AY178601S3 Gorilla g
35	111	8.8	111	9	AY178583S6	AY178583S6 Pan trogl
36	111	8.8	111	9	AY178583S9	AY178583S9 Pan trogl
37	111	8.8	111	9	AY178592S9	AY178592S9 Pan panis
38	111	8.8	111	9	AY178601S6	AY178601S6 Gorilla g
39	111	8.8	111	9	AY178601S9	AY178601S9 Gorilla g
40	111	8.8	111	9	AY178610S6	AY178610S6 Pongo pyg
41	111	8.8	111	9	AY178610S9	AY178610S9 Pongo pyg
42	109	8.6	313	6	AX915011	AX915011 Sequence
43	109	8.6	313	6	BD050544	BD050544 Sequence
44	105	8.3	105	9	AY178583S8	AY178583S8 Pan trogl
45	105	8.3	105	9	AY178592S8	AY178592S8 Pan panis

ALIGNMENTS

RESULT 1 AF337549 1993 bp mRNA linear PRI 15-JUN-2004
LOCUS Homo sapiens EKN1 (EKN1) mRNA, complete cds.
DEFINITION AF337549
ACCESSION AF337549
VERSION AF337549.1 GI:18478647
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1993)
AUTHORS Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B.,
Lytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J.,
Hannula-Jouppi,K. and Kere,J.
TITLE A candidate gene for developmental dyslexia encodes a nuclear
tetrapeptide repeat domain protein dynamically regulated in
brain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
MEDLINE 22882828
PUBMED 12954984
REFERENCE 2 (bases 1 to 1993)
AUTHORS Taipale,M. and Kere,J.
TITLE A gene disrupted by translocation breakpoint in chromosome 15q21
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1993)
AUTHORS Taipale,M. and Kere,J.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2001) Finnish Genome Center, University of
Helsinki, Tukholmankatu 2, Helsinki 00014, Finland
FEATURES
source
1. 1993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21.2"
1. 1993
/gene="EKN1"
/note="wildtype gene occasionally disrupted by chromosomal
translocation breakpoint"
369. 1631
CDS


```
/gene="EKX1"
/codon_start=1
/product="EKX1"
/protein_id="AAL73230.1"
/db_xref="GI:18478648"
/translation="MPLQVSDYSNQOTKTAVFLSLPLKGVCRDVFCTENYLKVN
PPFLPEAPLAPIDDESSKAKIGNDTIVFTLYKKEAMWETLSVTGVDKEMMORIREK
SILOAQRAKEATEAKAAREDOKYALSVMWKIEBERKKIEMDKENERIKAHKALE
AWKEYQKAEQKIKQBEKLCQKQEKIKYKSLRNLSLAPKGRNSN
IFTEKLEDSIPAPRSVGSIKINFTPRVPTALRESQVABEEHMLKHQAEPARMND
IAELCDLKEBKPEWLDKGNKLFATENYLAALINLNNKPLIYLNRACH
LKLNHKALEDSSKALELMPVTDNANARMAHVRGTFACQLELYVEGLQDYEA
LKIDPSNKIVQIDAEKIRNVIGTELKS"
```

ORIGIN

```
Query Match      100.0%; Score 1263; DB 9; Length 1993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCTCTCAGGTTAGGATTACAGCTGGCAGCAGACGAACTGCGGTCTTCTGTCT 60
Db 369 ATGCCTCTCAGGTTAGGATTACAGCTGGCAGCAGACGAACTGCGGTCTTCTGTCT 428

Qy 61 CTGCCCTCAAGCGGTGTGCTGAGACACGACGCTTCTGCACGGAACACTCTG 120
Db 429 CTGCCCTCAAGCGGTGTGCTGAGACACGACGCTTCTGCACGGAACACTCTG 488

Qy 121 AAGTCAACTTTCCTCCATTTTATTTTGAAGCACTTCTTATGTCTCCATAGCAGTGA 180
Db 489 AAGTCAACTTTCCTCCATTTTATTTTGAAGCACTTCTTATGTCTCCATAGCAGTGA 548

Qy 181 AGCGAAGCAAGATTGGGAATGACACCATTTCTTCCCTTGTATATAAAGAGCG 240
Db 549 AGCGAAGCAAGATTGGGAATGACACCATTTCTTCCCTTGTATATAAAGAGCG 608

Qy 241 GCCATGTGGAGACCTTTCTGTGACGGTCTTCACAAAGAGATGATGCAAGAAATTAGA 300
Db 609 GCCATGTGGAGACCTTTCTGTGACGGTCTTCACAAAGAGATGATGCAAGAAATTAGA 668

Qy 301 GAAAAATCTATTTTACAGCAACAGAGAGACAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db 669 GAAAAATCTATTTTACAGCAACAGAGAGACAAAGAGCTACAGAAAGCAAAAGCTGCA 728

Qy 361 GCAAAGCGGGAAGATCAAAAATAGCCTAAGTGTCTGATGATGATGATGATGATGATGAT 420
Db 729 GCAAAGCGGGAAGATCAAAAATAGCCTAAGTGTCTGATGATGATGATGATGATGATGAT 788

Qy 421 AGGAAAAAATAGAGATATGAAAGAAATGAAACGATGATGATGATGATGATGATGATGAT 480
Db 789 AGGAAAAAATAGAGATATGAAAGAAATGAAACGATGATGATGATGATGATGATGATGAT 848

Qy 481 GCCTGGAAAGATATCAAGAAAAAGCTGAGAGCAAAAAAATTTTCAGAGAGAGAGAAA 540
Db 849 GCCTGGAAAGATATCAAGAAAAAGCTGAGAGCAAAAAAATTTTCAGAGAGAGAGAAA 908

Qy 541 TTATGTCAAAAAGAAAAAGCAAAATTAAAGAGAGAGAAAAAATAAATAAAGAGCTTT 600
Db 909 TTATGTCAAAAAGAAAAAGCAAAATTAAAGAGAGAGAAAAAATAAATAAAGAGCTTT 968

Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAGGAGAGAAATTCAGAAAAATATATT 660
Db 969 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAGGAGAGAAATTCAGAAAAATATATT 1028

Qy 661 ACTGAGAGAGTTAAAGGAGAGACAGTATTCCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 720
Db 1029 ACTGAGAGAGTTAAAGGAGAGACAGTATTCCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 1088

Qy 721 AACTTTACCCCTCGAGTATTCCTCAACAGCTCTTCTGTAATCAAGTATGAGAGAGAG 780
Db 1089 AACTTTACCCCTCGAGTATTCCTCAACAGCTCTTCTGTAATCAAGTATGAGAGAGAG 1148

Qy 781 GAGTGGCTACACAAAGAGCTGAGCGCAAGAGCAATGAATCTGACATAGCTGAATCTT 840
Db 1148 GAGTGGCTACACAAAGAGCTGAGCGCAAGAGCAATGAATCTGACATAGCTGAATCTT 1200
```

```
Db 1149 GAGTGGCTACACAAAGAGCTGAGCGCAAGAGCAATGAATCTGACATAGCTGAATCT 1208
Qy 841 TCGGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 1209 TCGGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1268
Qy 901 TTTTGAACGAGAAACTATTTTGGCAGCTATCAATGCAATATATATTTAGCCATAAGACTAAAT 960
Db 1269 TTTTGAACGAGAAACTATTTTGGCAGCTATCAATGCAATATATATTTAGCCATAAGACTAAAT 1328
Qy 961 AATAAGATGCACTATTTGTAATTTGAACCGGCTGCTTGGCCACTTAAACCTTAAACCTTA 1020
Db 1329 AATAAGATGCACTATTTGTAATTTGAACCGGCTGCTTGGCCACTTAAACCTTAAACCTTA 1388
Qy 1021 CACAAGGCTATTGAAGATTTCTTCTAAGGCACTGGAATTTATGATGCCACTGTACAGAC 1080
Db 1389 CACAAGGCTATTGAAGATTTCTTCTAAGGCACTGGAATTTATGATGCCACTGTGTACAGAC 1448
Qy 1081 AATGCTAATCAAGAAATGAAGGACATGTACGACGTGGAAACAGCATTTCTGTCAACTAGAA 1140
Db 1449 AATGCTAATCAAGAAATGAAGGACATGTACGACGTGGAAACAGCATTTCTGTCAACTAGAA 1508
Qy 1141 TTGTATGTAGAAGCCCTACAGGATTTAAGACGGCACTTAAGATTTGATTCATCAACAAA 1200
Db 1509 TTGTATGTAGAAGCCCTACAGGATTTAAGACGGCACTTAAGATTTGATTCATCAACAAA 1568
Qy 1201 ATTGTACAAATTCATGCTGAGAGATTCGGAATTTAATTTCAAGGACAGAACTAAATCT 1260
Db 1569 ATTGTACAAATTCATGCTGAGAGATTCGGAATTTAATTTCAAGGACAGAACTAAATCT 1628
Qy 1261 TAA 1263
Db 1629 TAA 1631

RESULT 2
AX833451 Sequence 575 from Patent EP1347046.
LOCUS AX833451 1559 bp DNA linear PAT 15-DEC-2003
DEFINITION AX833451
ACCESSION AX833451
VERSION AX833451.1 GI:39919586
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isegaki, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
Masuho, Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 575 24-SEP-2003;
Research Association for Biotechnology (JP)
FEATURES
source 1..1559
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 79.0%; Score 998; DB 6; Length 1559;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCCTCTTTCAGGTTAGGATTACAGCTGGCAGCAGACGAACTGCGGTCTTCTGTCT 60
Db 43 ATGCCTCTTTCAGGTTAGGATTACAGCTGGCAGCAGACGAACTGCGGTCTTCTGTCT 102

Qy 61 CTGCCCTCAAGCGGTGTGCTGAGACACGACGCTTCTGACGAGAACTATCTG 120
Db 103 CTGCCCTCAAGCGGTGTGCTGAGACACGACGCTTCTGACGAGAACTATCTG 162
```


QY 121- AAGTCAATTTCTCCATTTTATTTATTTGAGGCAATTTCTTTATGCTCCCATAGACGATGAG 180
 Db |||||
 QY 163 AAGGTCACATTTCTCCATTTTATTTATTTGAGGCAATTTCTTTATGCTCCCATAGACGATGAG 222
 Db |||||
 QY 181 AGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCCACCTTGTATATAAAAAAGAGCG 240
 Db |||||
 QY 223 AGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCCACCTTGTATATAAAAAAGAGCG 282
 Db |||||
 QY 241 GGCATGTGGAGACCTTTCTGTGACGGGTGTGTGCAAGAGAGATGATGCAAGAATTTAGA 300
 Db |||||
 QY 283 GGCATGTGGAGACCTTTCTGTGACGGGTGTGTGCAAGAGAGATGATGCAAGAATTTAGA 342
 Db |||||
 QY 301 GAAATATCTATTTTACAGCAACAGAGAGAGCAAGAGAGCTTACAGAGAGCAAGAGCTGCA 360
 Db |||||
 QY 343 GAAATATCTATTTTACAGCAACAGAGAGAGCAAGAGAGCTTACAGAGAGCAAGAGCTGCA 402
 Db |||||
 QY 361 GCAAGCGGAGAGATCAAAAATATCCACTAAGTGTTCATGATGAAGATTTGAAGAGAGAG 420
 Db |||||
 QY 403 GCAAGCGGAGAGATCAAAAATATCCACTAAGTGTTCATGATGAAGATTTGAAGAGAGAG 462
 Db |||||
 QY 421 AGGAAAAAATAGAGATATGAAGAAATGAACGGATAAAGCCATTAAGCATTTGAA 480
 Db |||||
 QY 463 AGGAAAAAATAGAGATATGAAGAAATGAACGGATAAAGCCATTAAGCATTTGAA 522
 Db |||||
 QY 481 GCCTGGAAGATATCAAGAAAGCTGAGGAGCAAAAAAATTTACAGAGAGAGAGAA 540
 Db |||||
 QY 523 GCCTGGAAGATATCAAGAAAGCTGAGGAGCAAAAAAATTTACAGAGAGAGAGAA 582
 Db |||||
 QY 541 TTATGTCAAAAGAAAGCAATTAAGAGAGCAAAAAAATTAATAATAAGAGCTTT 600
 Db |||||
 QY 583 TTATGTCAAAAGAAAGCAATTAAGAGAGCAAAAAAATTAATAATAAGAGCTTT 642
 Db |||||
 QY 601 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAAGGAGAGAAATTCAGAAAAATATTT 660
 Db |||||
 QY 643 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAAGGAGAGAAATTCAGAAAAATATTT 702
 Db |||||
 QY 661 ACTGAGAGTTTAAAGAGAGAGATTTCTGCTCTCTGCTCTGCTGTCAGTATTAAGATC 720
 Db |||||
 QY 703 ACTGAGAGTTTAAAGAGAGAGATTTCTGCTCTCTGCTCTGTCAGTATTAAGATC 762
 Db |||||
 QY 721 AACTTTACCTCGAGTATTTCCACAGCTCTTCTGTGATCAAGTACAGTACAGAGAGAG 780
 Db |||||
 QY 763 AACTTTACCTCGAGTATTTCCACAGCTCTTCTGTGATCAAGTACAGTACAGAGAGAG 822
 Db |||||
 QY 781 GAGTGGCTACACAAACAAGCTGAGCAGCAAGAGCAATGAATFACTGACATAGCTGAACCT 840
 Db |||||
 QY 823 GAGTGGCTACACAAACAAGCTGAGCAGCAAGAGCAATGAATFACTGACATAGCTGAACCT 882
 Db |||||
 QY 841 TGCATTTTAAAGAGAGAGAAAGAACCCAGATGGTTGAAGGATTAAGGAAACCAATTTG 900
 Db |||||
 QY 883 TGCATTTTAAAGAGAGAGAAAGAACCCAGATGGTTGAAGGATTAAGGAAACCAATTTG 942
 Db |||||
 QY 901 TTTGCAACGGAAGAACTATTTGCGAGCTATCAATGATATATTTAGCCATTAAGACTAAT 960
 Db |||||
 QY 943 TTTGCAACGGAAGAACTATTTGCGAGCTATCAATGATATATTTAGCCATTAAGACTAAT 1002
 Db |||||
 QY 961 AATAAGATGCCATTTATTTGTAACCGGCTGCTTCCACCTAAAACTAAAAAACTTA 1020
 Db |||||
 QY 1003 AATAAGATGCCATTTATTTGTAACCGGCTGCTTCCACCTAAAACTAAAAAACTTA 1062
 Db |||||
 QY 1021 CACAAGGCTATTGAAGATTTCTTAAGGC 1049
 Db |||||
 QY 1063 CACAAGGCTATTGAAGATTTCTTAAGGC 1091
 Db |||||

RESULT 3
 AK095201
 LOCUS Homo sapiens cDNA FLJ37882 fis, clone BRSTN2000536.
 DEFINITION
 ACCESSION
 AK095201
 VERSION
 AK095201.1 GI:21754405
 KEYWORDS
 oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Wakamatsu, A., Hayaishi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
 Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
 Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
 Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Kanda, K., Kamihara, K., Yokoi, T., Furuya, T.,
 Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsura, N., Sato, K.,
 Tanikawa, M., Yanazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
 Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
 Hiraoaka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
 Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
 Hara, H., Tanase, T., Nomura, Y., Toqiya, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
 Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
 Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Koniyama, M.,
 Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
 Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
 Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
 Nakajima, Y., Mizuno, T., Shigetani, K., Senba, T., Matsumura, K.,
 Noguchi, Y., Mitsu, T., Morinaga, M., Sasaki, M., Togashi, T.,
 Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
 Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
 Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.
 and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 1559)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (04-JUL-2002)

Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

Location/Qualifiers

1. .1559

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="BRSTN2000536"

/tissue_type="subthalamic nucleus"

/clone_lib="BRSTN2"

/note="cloning vector: pMB18SFL3"

43. .1173

/note="unnamed protein product"

/codon_start=1

/protein_id="BAC04498.1"
/db_xref="GI:21754406"
/translation="MPQVSDVSWQQTAKVFLSLPLKGVGVDRDTPFCNTYLVKNF
PPFLFEAPLADIDDESKAKI GNDITVFLYKKEAMWETLSVTGVDKEMMQRIREK
SILQOARAKEATEAKAAKREDOQYALSWMKI EEEERKKI EDMKENERIKATKALE
AWKEYORKEAEOKI OREBKLCQKEKQI KEERKKI KYKSLRNLSANLAPKGRNSN
IFTEKLKEDSI PAPRSVGSIKI NPTPRVFTALRESOVAEEBEWLHKQAEARRAMNTD
IAELCDLKEBEKPEWLKDKGNKLFPATENYLAALNAYNLAIRLNNKPLLYLNRAACH
LKLNHLKAIEDSKATRIKMRHLRLHPTKLYKLMURFGM"

ORIGIN

Query Match		79.0%; Score 998; DB 9; Length 1559;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 1048; Conservative		0; Mismatches 1; Indels 0; Gaps 0;
Qy	1	ATGCCTCTTCAGGTAGCGATTACAGCTGGCAGACGACGAGCTGGCGTCTTCTCTCT 60
Db	43	ATGCCTCTTCAGGTAGCGATTACAGCTGGCAGCAGCAGACGAGCTGGCGTCTTCTCTCT 102
Qy	61	CTGCCCTCAAGGCGTGTGGTCAGACGACGACGCTGTCTGACGGAAACTATCTG 120
Db	103	CTGCCCTCAAGGCGTGTGGTCAGACGACGACGCTGTCTGACGGAAACTATCTG 162
Qy	121	AAGGTCAACTTCTCCATTTTATTGAGCATTTCTTTATGCTCCCATAGACGATGAG 180
Db	163	AAGGTCAACTTCTCCATTTTATTGAGCATTTCTTTATGCTCCCATAGACGATGAG 222
Qy	181	AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCACTTGTATAAAAAAGAGCG 240
Db	223	AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCACTTGTATAAAAAAGAGCG 282
Qy	241	GCCATGTGGAGACCCCTTCTGTACGGTGTTCACAAAGAGATGATCAAGAAATTAGA 300
Db	283	GCCATGTGGAGACCCCTTCTGTACGGTGTTCACAAAGAGATGATGACAAAGAAATTAGA 342
Qy	301	GAATAAATCTATTTTACAGCAACAGAGAGACCAAGAGCTACAGAGCAAGCAAGCTGCA 360
Db	343	GAATAAATCTATTTTACAGCAACAGAGAGACCAAGAGCTACAGAGCAAGCAAGCTGCA 402
Qy	361	GCAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATGGAAGAAGAG 420
Db	403	GCAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATGGAAGAAGAG 462
Qy	421	AGGAAAAATATAGAGATATGAAGAAATGAAACGGATAAAGCCACTTAAGCATTCGAA 480
Db	463	AGGAAAAATATAGAGATATGAAGAAATGAAACGGATAAAGCCACTTAAGCATTCGAA 522
Qy	481	GCCTGGAAAGAAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 540
Db	523	GCCTGGAAAGAAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 582
Qy	541	TTATGTCAAAAGAAAAAGCAAAATTAAGAAGAGAGAAAAAATAAATAATAAGAGCTTT 600
Db	583	TTATGTCAAAAGAAAAAGCAAAATTAAGAAGAGAGAAAAAATAAATAATAAGAGCTTT 642
Qy	601	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATT 660
Db	643	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATT 702
Qy	661	ACTGAGAGTTTAAAGGAAGACAGTATTCCTCTCTCTGCTGTGGCAGTATTAAAAATC 720
Db	703	ACTGAGAGTTTAAAGGAAGACAGTATTCCTCTCTCTGCTGTGGCAGTATTAAAAATC 762
Qy	721	AACCTTTACCCCTCGAGTATTCCTCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGAG 780
Db	763	AACCTTTACCCCTCGAGTATTCCTCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGAG 822
Qy	781	GAGTGGCTACACAAACAGCTGAGCAGAGAGCAATGAATCTCATAGCTGAACCTT 840
Db	823	GAGTGGCTACACAAACAGCTGAGCAGAGAGCAATGAATCTCATAGCTGAACCTT 882
Qy	841	TGCCATTTAAAGGAAGAGAAAAAGAACCCAGATGTTGTAAGGATTAAGGAGAAACAAATTG 900

Db	883	TGCGATTTAAAGAGAGAGAAAGAACCCAGATGGTTGAAGGATAAGGAGAAACAAATTG 942
Qy	901	TTTGCACGAGAAACTATTTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960
Db	943	TTTGCACGAGAAACTATTTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 1002
Qy	961	AATAAGATGCCATCTATTGTATTGAAACCGGCGTCTTGCACCTTAAAAAATACTTAA 1020
Db	1003	AATAAGATGCCATCTATTGTATTGAAACCGGCGTCTTGCACCTTAAAAAATACTTAA 1062
Qy	1021	CACAAGGCTATTGAAGATTTCTTCAAGGC 1049
Db	1063	CACAAGGCTATTGAAGATTTCTTCAAGGC 1091
RESULT 4		
BC062564		
LOCUS		
DEFINITION	Homo sapiens dyslexia susceptibility 1 candidate 1, mRNA (cdna	linear PRI 28-NOV-2003
ACCESSION	BC062564	
VERSION	BC062564.1	GI:38565951
KEYWORDS	MGC.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 1468)	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.B., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.D. and Marra, M.A., 2003	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE	22388257	
PUBMED	12477932	
REFERENCE	2 (bases 1 to 1468)	
AUTHORS	Strausberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK		
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgaps-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nih.gov Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,	


```
Qy 147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGACGAGCAAAAGATTGGGAATGA 206
Db 203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGACGAGCAAAAGATTGGGAATGA 262
Qy 207 CACCATTTGTTCTTACCTTGTATATAAAAGAACGGCCCATGTGGAGACCCCTTTCTGTGAC 266
Db 263 CACCATTTGTTCTTACCTTGTATATAAAAGAACGGCCCATGTGGAGACCCCTTTCTGTGAC 322
Qy 267 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 326
Db 323 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 382
Qy 327 GAGAGCAAAAGAGCTACAGAACCAAAAGCTGCAGCAAAAGCGGGAAGATCAAAAATACGC 386
Db 383 GAGAGCAAAAGAGCTACAGAACCAAAAGCTGCAGCAAAAGCGGGAAGATCAAAAATACGC 442
Qy 387 ACTAAGTGTCTATGATGAAGATTGAAGAGAGAGAGAGGAAAAAATAG 433
Db 443 ACTAAGTGTCTATGATGAAGATTGAAGAGAGAGAGAGGAAAAAATAG 489

RESULT 6
AX970431
LOCUS AX970431 489 bp DNA linear PAT 15-JAN-2004
DEFINITION Sequence 1234 from Patent EP1104808.
ACCESSION AX970431
VERSION AX970431.1 GI:40977781
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE ESTs and encoded human proteins
JOURNAL Patent: EP 1104808-A 1234 06-JUN-2001;
Genset (FR)
FEATURES
source
1..489
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
57..>488
/note="unamed protein product"
/codon_start=1
/protein_id="CAF15539.1"
/db_xref="GI:40977782"
/translation="MPLQVSDYKXWQTKTAVFLSLPLKGVCRDIDVFCTENYLKVPF
PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIEK
SILQAEKAEATEAKAAKREDQKVALSVNMKIEEBERKKI"
CDS
32..322
/codon_start=1
/protein_id="CAF15539.1"
/db_xref="GI:40977782"
/translation="MPLQVSDYKXWQTKTAVFLSLPLKGVCRDIDVFCTENYLKVPF
PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIEK
SILQAEKAEATEAKAAKREDQKVALSVNMKIEEBERKKI"
ORIGIN
Query Match 32.2%; Score 407; DB 6; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 27 CTGGCAGCAGACGAGACTGCGGTCTTTCTGCTCTGCCCCCTCAAAGCGGTGCGTCAG 86
Db 83 CTGGCAGCAGACGAGACTGCGGTCTTTCTGCTCTGCCCCCTCAAAGCGGTGCGTCAG 142
Qy 87 AGACACGACGCTGTTCTGCACGGAAACTATCTGAAGTCAACTTCTCCATTTTATT 146
Db 143 AGACACGACGCTGTTCTGCACGGAAACTATCTGAAGTCAACTTCTCCATTTTATT 202
Qy 147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGACGAGCAAAAGATTGGGAATGA 206
Db 203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGACGAGCAAAAGATTGGGAATGA 262
Qy 207 CACCATTTGTTCTTACCTTGTATATAAAAGAACGGCCCATGTGGAGACCCCTTTCTGTGAC 266
Db 263 CACCATTTGTTCTTACCTTGTATATAAAAGAACGGCCCATGTGGAGACCCCTTTCTGTGAC 322
Qy 267 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 326
Db 323 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 382
```

```
Db 323 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 382
Qy 327 GAGAGCAAAAGAGCTACAGAACCAAAAGCTGCAGCAAAAGCGGGAAGATCAAAAATACGC 386
Db 383 GAGAGCAAAAGAGCTACAGAACCAAAAGCTGCAGCAAAAGCGGGAAGATCAAAAATACGC 442
Qy 387 ACTAAGTGTCTATGATGAAGATTGAAGAGAGAGAGGAAAAAATAG 433
Db 443 ACTAAGTGTCTATGATGAAGATTGAAGAGAGAGAGGAAAAAATAG 489

RESULT 7
BD109150
LOCUS BD109150 489 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD109150
VERSION BD109150.1 GI:23203968
KEYWORDS JP 2002010789-A/1227.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 1227 15-JAN-2002;
GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/1227
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI
GIORDANO
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,PC
C12N1/21,
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00,PC
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS 57..488.
FEATURES
source
1..489
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 32.2%; Score 407; DB 6; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 27 CTGGCAGCAGACGAGACTGCGGTCTTTCTGCTCTGCCCCCTCAAAGCGGTGCGTCAG 86
Db 83 CTGGCAGCAGACGAGACTGCGGTCTTTCTGCTCTGCCCCCTCAAAGCGGTGCGTCAG 142
Qy 87 AGACACGACGCTGTTCTGCACGGAAACTATCTGAAGTCAACTTCTCCATTTTATT 146
Db 143 AGACACGACGCTGTTCTGCACGGAAACTATCTGAAGTCAACTTCTCCATTTTATT 202
Qy 147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGACGAGCAAAAGATTGGGAATGA 206
Db 203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGACGAGCAAAAGATTGGGAATGA 262
Qy 207 CACCATTTGTTCTTACCTTGTATATAAAAGAACGGCCCATGTGGAGACCCCTTTCTGTGAC 266
Db 263 CACCATTTGTTCTTACCTTGTATATAAAAGAACGGCCCATGTGGAGACCCCTTTCTGTGAC 322
Qy 267 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 326
Db 323 GGGTGTTCACAAAGAGATGATGCAAGAAATAGAGAAAAATCTATTTTACAAGCACAGA 382
```



```

Db      229 AGCAGCAAGCAAGATTGGGAATGACACCATTC 262
|||||
AC021668      43886 bp      DNA      linear      HTG 29-MAR-2001
Homo sapiens chromosome 15 clone RP11-13306 map 15, LOW-PASS
SEQUENCE SAMPLING.
AC021668      4 GI:13488020
HTG; HTGS PHASE0.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 43886)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguelavkiy,L., Boukhealter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,X., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivat,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Mar 29, 2001 this sequence version replaced gi:11612359.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WTHR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L1276
            Center clone name: 133_O_6
            -----
            * NOTE: This record contains 53 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *
            * 1      745: contig of 745 bp in length
            *      746      845: gap of 100 bp
            *      846      1576: contig of 731 bp in length
            *      1577      1676: gap of 100 bp
            *      1677      2410: contig of 734 bp in length
            *      2411      2510: gap of 100 bp
            *      2511      3232: contig of 722 bp in length
            *      3233      3332: gap of 100 bp
            *
            * 3333      4060: contig of 728 bp in length
            * 4061      4160: gap of 100 bp
            * 4161      4891: contig of 731 bp in length
            * 4892      4991: gap of 100 bp
            * 4992      5708: contig of 717 bp in length
            * 5709      5808: gap of 100 bp
            * 5809      6337: contig of 729 bp in length
            * 6338      6537: gap of 100 bp
            * 6539      7344: contig of 707 bp in length
            * 7345      7444: gap of 100 bp
            * 7445      8167: contig of 723 bp in length
            * 8168      8267: gap of 100 bp
            * 8268      8995: contig of 728 bp in length
            * 8996      9095: gap of 100 bp
            * 9096      9839: contig of 744 bp in length
            * 9840      9939: gap of 100 bp
            * 9940      10653: contig of 714 bp in length
            * 10654      10753: gap of 100 bp
            * 10754      11497: contig of 744 bp in length
            * 11498      11597: gap of 100 bp
            * 11598      12315: contig of 718 bp in length
            * 12316      12415: gap of 100 bp
            * 12416      13136: contig of 721 bp in length
            * 13137      13236: gap of 100 bp
            * 13237      13967: contig of 731 bp in length
            * 13968      14067: gap of 100 bp
            * 14068      14802: contig of 735 bp in length
            * 14803      14902: gap of 100 bp
            * 14903      15642: contig of 740 bp in length
            * 15643      15742: gap of 100 bp
            * 15743      16477: contig of 735 bp in length
            * 16478      16577: gap of 100 bp
            * 16578      17301: contig of 724 bp in length
            * 17302      17401: gap of 100 bp
            * 17401      18129: contig of 728 bp in length
            * 18130      18229: gap of 100 bp
            * 18230      18958: contig of 729 bp in length
            * 18959      19058: gap of 100 bp
            * 19059      19787: contig of 729 bp in length
            * 19788      19887: gap of 100 bp
            * 19888      20627: contig of 740 bp in length
            * 20628      20727: gap of 100 bp
            * 20728      21449: contig of 722 bp in length
            * 21450      21549: gap of 100 bp
            * 21550      22275: contig of 726 bp in length
            * 22276      22375: gap of 100 bp
            * 22376      23103: contig of 728 bp in length
            * 23104      23203: gap of 100 bp
            * 23204      23939: contig of 736 bp in length
            * 23940      24039: gap of 100 bp
            * 24040      24775: contig of 736 bp in length
            * 24776      24875: gap of 100 bp
            * 24876      25602: contig of 727 bp in length
            * 25603      25702: gap of 100 bp
            * 25703      26518: contig of 716 bp in length
            * 26519      26518: gap of 100 bp
            * 26519      27260: contig of 742 bp in length
            * 27261      27360: gap of 100 bp
            * 27361      28107: contig of 747 bp in length
            * 28108      28207: gap of 100 bp
            * 28208      28931: contig of 724 bp in length
            * 28932      29031: gap of 100 bp
            * 29032      29776: contig of 745 bp in length
            * 29777      29876: gap of 100 bp
            * 29877      30592: contig of 716 bp in length
            * 30593      30692: gap of 100 bp
            * 30693      31422: contig of 730 bp in length
            * 31423      31522: gap of 100 bp
            * 31523      32255: contig of 733 bp in length
            * 32256      32355: gap of 100 bp
            * 32356      33073: contig of 718 bp in length
            * 33074      33173: gap of 100 bp
            * 33174      33891: contig of 718 bp in length

```



```

Oy 584 TAAATATAGAGCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGG 638
Db 134630 TAAATATAGAGCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGG 134576

RESULT 12 102(b)
AC016527/c
LOCUS AC016527 208181 bp DNA linear HTG 04-JUN-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-460A24 map 15q21, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC016527
VERSION AC016527.2 GI:8247792
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 208181)
Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
Sequencing of human chromosome 15 DI5S146-D15S117 region
Unpublished
2 (bases 1 to 208181)
Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
and Hood,L.
Direct Submission
Submitted (02-DSC-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Jun 4, 2000 this sequence version replaced gi:6503271.
----- Genome Center
Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowen@systemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
-----
* NOTE: This record contains 191 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
* 857 856: contig of 856 bp in length
* 957 956: gap of unknown length
* 1817 1817: contig of 861 bp in length
* 1818 1817: gap of unknown length
* 1918 2890: contig of 973 bp in length
* 2891 2990: gap of unknown length
* 2991 3851: contig of 861 bp in length
* 3852 3951: gap of unknown length
* 3952 5018: contig of 1067 bp in length
* 5019 5118: gap of unknown length
* 5119 5991: contig of 873 bp in length
* 5992 6091: gap of unknown length
* 6092 7038: contig of 947 bp in length
* 7039 7138: gap of unknown length
* 8022 8121: contig of 883 bp in length
* 8122 9152: contig of 1031 bp in length
* 9153 9252: gap of unknown length

```

```

* 9253 10196: contig of 944 bp in length
* 10197 10296: gap of unknown length
* 10297 11211: contig of 915 bp in length
* 11212 11311: gap of unknown length
* 11312 12220: contig of 909 bp in length
* 12321 12320: gap of unknown length
* 12321 13554: contig of 1234 bp in length
* 13555 13654: gap of unknown length
* 13655 14552: contig of 898 bp in length
* 14553 14652: gap of unknown length
* 14653 15827: contig of 1175 bp in length
* 15828 15927: gap of unknown length
* 15928 16959: contig of 1032 bp in length
* 16960 17059: gap of unknown length
* 17060 18020: contig of 961 bp in length
* 18021 18120: gap of unknown length
* 18121 19061: contig of 941 bp in length
* 19062 19161: gap of unknown length
* 19162 20206: contig of 1045 bp in length
* 20207 20306: gap of unknown length
* 20307 21198: contig of 892 bp in length
* 21199 21298: gap of unknown length
* 21299 22311: contig of 1013 bp in length
* 22312 23315: contig of 904 bp in length
* 23316 23415: gap of unknown length
* 23416 24425: contig of 1010 bp in length
* 24426 24525: gap of unknown length
* 24526 25448: contig of 923 bp in length
* 25449 25548: gap of unknown length
* 25549 26468: contig of 920 bp in length
* 26469 26589: gap of unknown length
* 26590 27410: contig of 842 bp in length
* 27411 27510: gap of unknown length
* 27511 28462: contig of 952 bp in length
* 28463 28562: gap of unknown length
* 28563 29473: contig of 911 bp in length
* 29474 29573: gap of unknown length
* 29574 30791: contig of 1218 bp in length
* 30792 30891: gap of unknown length
* 30892 32425: contig of 1534 bp in length
* 32426 32526: gap of unknown length
* 32526 33459: contig of 934 bp in length
* 33460 33559: gap of unknown length
* 33560 34491: contig of 932 bp in length
* 34492 34591: gap of unknown length
* 34592 35495: contig of 904 bp in length
* 35496 35595: gap of unknown length
* 35596 36461: contig of 865 bp in length
* 36461 37496: gap of unknown length
* 37497 37596: contig of 936 bp in length
* 37597 38516: gap of unknown length
* 38517 38616: gap of unknown length
* 38617 39598: contig of 982 bp in length
* 39599 39699: gap of unknown length
* 39700 40626: contig of 927 bp in length
* 40627 40725: gap of unknown length
* 40726 41999: contig of 1274 bp in length
* 42000 42099: gap of unknown length
* 42100 43057: contig of 958 bp in length
* 43058 43157: gap of unknown length
* 43158 44092: contig of 935 bp in length
* 44093 44192: gap of unknown length
* 44193 45076: contig of 884 bp in length
* 45077 45176: gap of unknown length
* 45177 46066: contig of 890 bp in length
* 46067 46166: gap of unknown length
* 46167 47021: contig of 855 bp in length
* 47022 47122: gap of unknown length
* 47123 48056: contig of 935 bp in length
* 48057 48156: gap of unknown length
* 48157 49117: contig of 961 bp in length

```



```
* 49118 49217: gap of unknown length
* 49218 50340: contig of 1023 bp in length
* 50341 51226: contig of 886 bp in length
* 51227 51326: gap of unknown length
* 51327 52251: contig of 925 bp in length
* 52252 52351: gap of unknown length
* 52352 53222: contig of 871 bp in length
* 53223 53322: gap of unknown length
* 53323 54237: contig of 915 bp in length
* 54238 54337: gap of unknown length
* 54338 55323: contig of 986 bp in length
* 55324 55423: gap of unknown length
* 55424 56377: contig of 954 bp in length
* 56378 56477: gap of unknown length
* 56478 57356: contig of 879 bp in length
* 57357 58363: contig of 907 bp in length
* 58364 58463: gap of unknown length
* 58464 59359: contig of 896 bp in length
* 59360 59459: gap of unknown length
* 59460 60347: contig of 888 bp in length
* 60348 60447: gap of unknown length
* 60448 61382: contig of 935 bp in length
* 61383 61482: gap of unknown length
* 61483 62469: contig of 987 bp in length
* 62470 62569: gap of unknown length
* 62570 63425: contig of 856 bp in length
* 63426 63525: gap of unknown length
* 63526 64491: contig of 966 bp in length
* 64492 64591: gap of unknown length
* 64592 65503: contig of 912 bp in length
* 65504 65603: gap of unknown length
* 65604 66496: contig of 893 bp in length
* 66497 66596: gap of unknown length
* 66597 67454: contig of 858 bp in length
* 67455 67554: gap of unknown length
* 67555 68608: contig of 1054 bp in length
* 68609 68708: gap of unknown length
* 68709 69566: contig of 858 bp in length
* 69567 69666: gap of unknown length
* 69667 70597: contig of 931 bp in length
* 70598 70697: gap of unknown length
* 70698 71587: contig of 890 bp in length
* 71588 71687: gap of unknown length
* 71689 72595: contig of 908 bp in length
* 72596 72695: gap of unknown length
* 72696 73659: contig of 974 bp in length
* 73670 73769: gap of unknown length
* 73770 74685: contig of 916 bp in length
* 74686 74785: gap of unknown length
* 74786 75768: contig of 983 bp in length
* 75769 75868: gap of unknown length
* 75869 76786: contig of 918 bp in length
* 76787 76886: gap of unknown length
* 76887 77743: contig of 857 bp in length
* 77744 77843: gap of unknown length
* 77844 79065: contig of 1222 bp in length
* 79066 79165: gap of unknown length
* 79166 80841: contig of 1676 bp in length
* 80842 80942 81875: contig of 934 bp in length

Query Match      11.9%; Score 150; DB 2; Length 208181;
Best Local Similarity 100.0%; Pred. No. 2.9e-65;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 AGGTCAACTTCTCCATTTTATTTAGGACATTTCTTTATGCTCCCATAGACGATGAGA 181
Db 22658 AGGTCAACTTCTCCATTTTATTTAGGACATTTCTTTATGCTCCCATAGACGATGAGA 22599
QY 182 GCAGCAAGCAAGATTGGGAATGACACCATTTGTTTACCTTGTATATAAAGAGCGG 241
|||||
```

```
Db 22598 GCAGCAAGCAAGATTGGGAATGACACCATTTGTTTACCTTGTATATAAAGAGCGG 22539
QY 242 CCATGTGGGAGACCTTTCTGTGACGGGTG 271
|||||
Db 22538 CCATGTGGGAGACCTTTCTGTGACGGGTG 22509
|||||

RESULT 13
LOCUS AY178583S2 147 bp DNA linear PRI 02-OCT-2003
DEFINITION Pan troglodytes EKN1 (EKN1) gene, exon 3.
ACCESSION AY178584
VERSION AY178584.1 GI:27804579
KEYWORDS
SEGMENT
SOURCE 2 of 9
ORGANISM Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 147)
AUTHORS Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B.,
Lyytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J.,
Hannula-Jouppi,K. and Kere,J.
TITLE A candidate gene for developmental dyslexia encodes a nuclear
tetrapeptide repeat domain protein dynamically regulated in
brain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
MEDLINE 22882828
PUBMED 12954984
REFERENCE 2 (bases 1 to 147)
AUTHORS Kaminen,N. and Kere,J.
TITLE EKN1 gene orthologs in chimpanzee, pygmy chimpanzee, gorilla and
orangutan
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 147)
AUTHORS Kaminen,N. and Kere,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2002) Department of Medical Genetics, University
of Helsinki, PO Box 63, Haartmaninkatu 8, Helsinki FIN-00014,
Finland

FEATURES
source Location/Qualifiers
1..147
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
1..147
/gene="EKN1"
/number=3

exon

ORIGIN
Query Match      11.6%; Score 147; DB 9; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.1e-63;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GTCAACTTCTCCATTTTATTTAGGACATTTCTTTATGCTCCCATAGACGATGAGC 183
Db 1 GTCAACTTCTCCATTTTATTTAGGACATTTCTTTATGCTCCCATAGACGATGAGC 60
|||||
QY 184 AGCAAGCAAGATTGGGAATGACACCATTTGTTTACCTTGTATATAAAGAGCGG 243
Db 61 AGCAAGCAAGATTGGGAATGACACCATTTGTTTACCTTGTATATAAAGAGCGG 120
|||||
QY 244 ATGTGGGAGACCTTTCTGTGACGGGT 270
Db 121 ATGTGGGAGACCTTTCTGTGACGGGT 147
|||||

RESULT 14
LOCUS AY178583S5 147 bp DNA linear PRI 02-OCT-2003
DEFINITION Pan troglodytes EKN1 (EKN1) gene, exon 6.
ACCESSION AY178587
VERSION AY178587.1 GI:27804582
```


Qy 408 TGAAGAAGAGAGAGAAAAA 429
Db 486 TGAAGAAGAGAGAGAAAAA 507

RESULT 5

US-09-854-133-30
; Sequence 30, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 1316
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-30

Query Match 1.7%; Score 22; DB 4; Length 1316;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 408 TGAAGAAGAGAGAGAAAAA 429
Db 486 TGAAGAAGAGAGAGAAAAA 507

RESULT 6

US-09-248-796A-9280
; Sequence 9280, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1999-02-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 9280
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-9280

Query Match 1.6%; Score 20; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 564 TAAAGAAGAGAGAAAAA 583
Db 113 TAAAGAAGAGAGAAAAA 132

RESULT 7

US-09-949-016-17196
; Sequence 17196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17196
; LENGTH: 118143
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17196

Query Match 1.6%; Score 20; DB 4; Length 118143;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 644 ATTCAGAAATATATTTACT 663
Db 9735 ATTCAGAAATATATTTACT 9754

RESULT 8

US-09-949-016-16513/c
; Sequence 16513, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16513
; LENGTH: 177293
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16513

Query Match 1.6%; Score 20; DB 4; Length 177293;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 548 AAAAAGAAAGCAATTTAA 567
Db 44561 AAAAAGAAAGCAATTTAA 44542

RESULT 9

US-09-248-796A-12670
; Sequence 12670, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074,725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 12670
;; LENGTH: 216
;; TYPE: DNA
;; ORGANISM: Candida albicans
US-09-248-796A-12670

Query Match 1.5%; Score 19; DB 4; Length 216;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 572 GAAGAAAAAAATAATAA 590
Db 63 GAAGAAAAAAATAATAA 81

RESULT 10
US-09-513-999C-11853/c
; Sequence 11853 Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 11853
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 241
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 242
; OTHER INFORMATION: w=a or t
US-09-513-999C-11853

Query Match 1.5%; Score 19; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 570 AGGAAGAAAAATAATAA 588
Db 180 AGGAAGAAAAATAATAA 162

RESULT 11
US-08-343-443B-45/c
; Sequence 45, Application US/08343443B
; Patent No. 5968734
; GENERAL INFORMATION:
; APPLICANT: Aurias, Alain
; APPLICANT: Delattre, Olivier
; APPLICANT: Desmaze, Chantal
; APPLICANT: Melot, Thomas
; APPLICANT: Peter, Martine
; APPLICANT: Ploougaestel, Beatrice
; APPLICANT: Thomas, Gilles

;; APPLICANT: Zucman, Jessica
;; TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
;; TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
;; TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
;; TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
;; TITLE OF INVENTION: TRANSLATIONS
;; NUMBER OF SEQUENCES: 129
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Weiser & Associates
;; STREET: 230 South Fifteenth Street
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: AEDIT 1.0 DOS text editor
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/343,443B
;; FILING DATE: 18-NOV-1994
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR93/00494
;; FILING DATE: 19-MAY-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 92/06123
;; FILING DATE: 20-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Weiser, Gerard J.
;; REGISTRATION NUMBER: 19,763
;; REFERENCE/DOCKET NUMBER: 989.6121P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-875-8383
;; TELEFAX: 215-875-8394
;; INFORMATION FOR SEQ ID NO: 45:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 425 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-343-443B-45

Query Match 1.5%; Score 19; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 565 AAAGAAGGAAGAAAAAAA 583
Db 275 AAAGAAGGAAGAAAAAAA 257

RESULT 12
US-09-540-236-389
; Sequence 389, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 389
; LENGTH: 507
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-389
Query Match 1.5%; Score 19; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 54;

Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY 848 TAAAGAAGAGAAAGAA 866
|||||
Db 383 TAAAGAAGAGAAAGAA 401

RESULT 13

```

RES001 13
US-09-949-016-23803
; Sequence 23803, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23803
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-23803

```

Qy 290 AAAGAATTAGAGAAAAATC 308
|||
Db 180 AAAGAATTAGAGAAAAATC 198

RESIT.T 14

```

RESUL 14
US-09-949-016-23804
; Sequence 23804, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23804
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-23804

```

Qy 290 AAAGAATTAGAGAAAAATC 308
|||||
Db 422 AAAGAATTAGAGAAAAATC 440

RESIT.T 15

```

RES001.13
US-09-949-016-51133/c
; Sequence 51133, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51133
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-51133

```

Qy 550 AAAGAAAGCAAATTAAAG 568
|||||
Db 107 AAAGAAAGCAAATTAAAG 89

Search completed: July 8, 2005, 12:07:45
Job time : 248 secs

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:36:52 ; Search time 5659 Seconds
(without alignments)
10814.444 Million cell updates/sec

Title: US-10-681-199-1

Perfect score: 1263

Sequence: 1 atgctcttcaggttagcga.....gaacagaactaaatcttaa 1263

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1263	100.0	1993	9 AF337549	AF337549 Homo sapi
2	1047.4	82.9	1559	6 AX833451	AX833451 Sequence
3	1047.4	82.9	1559	9 AK095201	AK095201 Homo sapi
4	1047.2	82.9	1468	9 BC062564	BC062564 Homo sapi
5	846	67.0	1946	10 BC028462	BC028462 Mus muscu
6	485.2	38.4	2045	5 BC077575	BC077575 Xenopus l
7	432.6	34.3	489	6 AR413597	AR413597 Sequence
8	432.6	34.3	489	6 AX970431	AX970431 Sequence
9	432.6	34.3	489	6 BD109150	BD109150 EST and e
10	401.6	31.8	464	6 C0685075	C0685075 Sequence
11	390.2	30.9	1443	5 BC065881	BC065881 Danio rer
12	298	23.6	761	5 CR387522	CR387522 Gallus ga
13	275	21.8	1133	5 CR389904	CR389904 Gallus ga
14	241	19.1	114149	9 AC022083	AC022083 Homo sapi
15	239.4	19.0	43886	2 AC021668	AC021668 Homo sapi
16	239.4	19.0	151133	9 AC013355	AC013355 Homo sapi
17	234.6	18.6	208181	2 AC016527	AC016527 Homo sapi
18	227.8	18.0	231	9 AY178586	AY178586 Pan trogl
19	227.8	18.0	231	9 AY178595	AY178595 Pan panis

20	223	17.7	231	9	AY178601S4	AY178604 Gorilla g
21	219.8	17.4	231	9	AY178610S4	AY178613 Pongo pyg
22	178.8	14.2	313	6	AX915011	AX915011 Sequence
23	178.8	14.2	313	6	BD050544	BD050544 Sequence
24	162.8	12.9	164	6	AX918710	AX918710 Sequence
25	162.8	12.9	164	6	BD054243	BD054243 Sequence
26	151.4	12.0	153	9	AY178583S7	AY178589 Pan trogl
27	151.4	12.0	153	9	AY178592S7	AY178598 Pan panis
28	151.4	12.0	153	9	AY178601S7	AY178607 Gorilla g
29	151.4	12.0	153	9	AY178610S7	AY178616 Pongo pyg
30	147	11.6	147	9	AY178583S2	AY178584 Pan trogl
31	147	11.6	147	9	AY178583S5	AY178587 Pan trogl
32	147	11.6	147	9	AY178592S2	AY178593 Pan panis
33	147	11.6	147	9	AY178592S5	AY178596 Pan panis
34	147	11.6	147	9	AY178601S2	AY178602 Gorilla g
35	147	11.6	147	9	AY178601S5	AY178605 Gorilla g
36	147	11.6	147	9	AY178610S5	AY178611 Pongo pyg
37	145.4	11.5	147	9	AY178610S5	AY178614 Pongo pyg
38	134.4	10.6	1187	6	CQ732459	CQ732459 Sequence
39	133.4	10.6	135	9	AY178583S3	AY178585 Pan trogl
40	133.4	10.6	135	9	AY178592S3	AY178594 Pan panis
41	133.4	10.6	135	9	AY178601S3	AY178603 Gorilla g
42	131.8	10.4	135	9	AY178610S3	AY178612 Pongo pyg
43	125.4	9.9	187880	2	AC114669	AC114669 Mus muscu
44	121.4	9.6	123	9	AY178592S1	AY178592 Pan panis
45	121.4	9.6	123	9	AY178610S1	AY178610 Pongo pyg

ALIGNMENTS

RESULT 1	AF337549	1993 bp	mRNA	linear	PRI 15-JUN-2004
LOCUS	AF337549	Homo sapiens	EKN1 (EKN1)	mRNA, complete cds.	
DEFINITION	AF337549	Homo sapiens	EKN1 (EKN1)	mRNA, complete cds.	
ACCESSION	AF337549	Homo sapiens	EKN1 (EKN1)	mRNA, complete cds.	
VERSION	AF337549.1	GI:18478647			
KEYWORDS					
SOURCE		Homo sapiens	(human)		
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
		1 (bases 1 to 1993)			
		Taipale, M., Muller, K., Kaaranen, M., Lindsberg, P.J.,			
		Hannula-Jouppi, K. and Kere, J.			
TITLE		A candidate gene for developmental dyslexia encodes a nuclear			
		tetratricopeptide repeat domain protein dynamically regulated in			
		brain			
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)			
MEDLINE		22882828			
PUBMED		12954984			
REFERENCE		2 (bases 1 to 1993)			
AUTHORS		Taipale, M. and Kere, J.			
TITLE		A gene disrupted by translocation breakpoint in chromosome 15q21			
JOURNAL		Unpublished			
REFERENCE		3 (bases 1 to 1993)			
AUTHORS		Taipale, M. and Kere, J.			
TITLE		Direct Submission			
JOURNAL		Submitted (17-JAN-2001) Finnish Genome Center, University of			
		Heisinki, Tukholmankatu 2, Helsinki 00014, Finland			
FEATURES		Location/Qualifiers			
source		1..1993			
		/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/db_xref="taxon:9606"			
		/chromosome="15"			
		/map="15q21.2"			
		1..1993			
		/gene="EKN1"			
gene		/note="wildtype gene occasionally disrupted by chromosomal			
		translocation breakpoint"			
		369..1631			
CDS					

```
/gene="EKNI"
/codon_start=1
/product="EKNI"
/protein_id="AAL73230.1"
/db_xref="GI:18478648"
/translation="MPLQVSDYSGNQQTAVFLSLPLKGVGVDRDVFCTENYLKVN
PPFLFEALVAPIDDESSKAKIGNDTIVFTLYKKEAMWETLSVTGVDKEMMORIREK
SILQAEKAEATEAKAAKREOKYALSVMWKIEERKKIEMKENERIKATKALE
AWQEYQKABEQKKIQRKELCKQKEKKIKYKSLUTRNLSLAPKGRNSN
IFTEKLEDSIPAPRSVGSIKINFTPRVFTALRESQVAEEBEEHLKHQAEARRMND
IAELCDLKEBEKPEWLKGNKLFATENYLAANIINFLRNKMLLLYLNRACH
LKLNLHKAIEDSSKALELLMPPVTDNANRMAHVRRGTAFCQLELYVEGLQDYEA
LKIDPSNKIQVIDAEKIRNVIQGTLSK"
```

ORIGIN

```
Query Match      100.0%; Score 1263; DB 9; Length 1993;
Best Local Similarity 100.0%; Pred. No. 3e-237;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1  ATGCCTCTTCAGGTAGGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTCTGTCT 60
Db   369 ATGCCTCTTCAGGTAGGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTCTGTCT 428

Qy   61  CTGCCCTCAAAGCGGTGTGCGTCAGAGACACGACGACGTTCCTGCACGGAAAACTATCTG 120
Db   429 CTGCCCTCAAAGCGGTGTGCGTCAGAGACACGACGACGTTCCTGCACGGAAAACTATCTG 488

Qy   121 AAGTCAACTTTCCTCCATTTTATTTTGAAGCAATTTCTTTATGCTCCCATAGAGATGAG 180
Db   489 AAGTCAACTTTCCTCCATTTTATTTTGAAGCAATTTCTTTATGCTCCCATAGAGATGAG 548

Qy   181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTCCCTTGTATATAAAGAGCG 240
Db   549 AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTCCCTTGTATATAAAGAGCG 608

Qy   241 GCCATGTGGAGACCCCTTCTGTGACCGGTGTCACAAAGAGATGATGCAAAAGATTAGA 300
Db   609 GCCATGTGGAGACCCCTTCTGTGACCGGTGTCACAAAGAGATGATGCAAAAGATTAGA 668

Qy   301 GAAAAATCTATTTTACAAGCAACAAGAGAGACAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db   669 GAAAAATCTATTTTACAAGCAACAAGAGAGACAAAGAGCTACAGAAAGCAAAAGCTGCA 728

Qy   361 GCAAGCGGNAGATCAAAATAGCCACTAAGTGTCTGATGATGATGATGATGATGATGAT 420
Db   729 GCAAGCGGNAGATCAAAATAGCCACTAAGTGTCTGATGATGATGATGATGATGATGAT 788

Qy   421 AGGAAAAAATAGAGATATGAAAGAAATGAAACGGATATAAAGCCACTTAAAGCAATTGGAA 480
Db   789 AGGAAAAAATAGAGATATGAAAGAAATGAAACGGATATAAAGCCACTTAAAGCAATTGGAA 848

Qy   481 GCCTGGAAGAAATATCAAGAAAAAGCTGAGAGCAAAAAAATTCAGAGAGAGAGAGAAA 540
Db   849 GCCTGGAAGAAATATCAAGAAAAAGCTGAGAGCAAAAAAATTCAGAGAGAGAGAGAAA 908

Qy   541 TTAATGTCAAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db   909 TTAATGTCAAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 968

Qy   601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAGAAATTCAGAAAAATATAT 660
Db   969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAGAAATTCAGAAAAATATAT 1028

Qy   661 ACTGAGAAGTTAAAGGAGACAGTATTCCTGCTCCTCGCTCTGTGGCAGTATTAAATC 720
Db   1029 ACTGAGAAGTTAAAGGAGACAGTATTCCTGCTCCTCGCTCTGTGGCAGTATTAAATC 1088

Qy   721 AACTTTACCCCTCGAGTATTCACACAGCTCTTCGTGATCAACAGTACAGAGAGAGAG 780
Db   1089 AACTTTACCCCTCGAGTATTCACACAGCTCTTCGTGATCAACAGTACAGAGAGAGAG 1148

Qy   781 GAGTGGCTTACACAAAGCTGAGCGACGAGAGCAATGAATACTGACATAGCTGAATCT 840
```

```
Db   1149 GAGTGGCTTACACAAAGCTGAGGCACGAGAGCAATGAATACTGACATAGCTGAACCT 1208
Qy   841 TGGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db   1209 TGGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1268
Qy   901 TTTGCAACGGAAGAACTATTTGGCAGCTATCAATGTCATATAATTTAGCCATAAGAGACTAAAT 960
Db   1269 TTTGCAACGGAAGAACTATTTGGCAGCTATCAATGTCATATAATTTAGCCATAAGAGACTAAAT 1328
Qy   961 AATAAGATGCCACTATTTGTAATTTGAACCGGGCTGCTGCGACCTTAAACCTTAAACCTTA 1020
Db   1329 AATAAGATGCCACTATTTGTAATTTGAACCGGGCTGCTGCGACCTTAAACCTTAAACCTTA 1388
Qy   1021 CACAGGCTATTTGAAGATTTCTTCTAAGGCATGGAATTTATTTGATGCCACCTGTTACAGAC 1080
Db   1389 CACAGGCTATTTGAAGATTTCTTCTAAGGCATGGAATTTATTTGATGCCACCTGTTACAGAC 1448
Qy   1081 AATGCTAATCAAGAAATGAAGGCACATGTACGACGTGGAACAGCATTTCTGTCAACTAGAA 1140
Db   1449 AATGCTAATCAAGAAATGAAGGCACATGTACGACGTGGAACAGCATTTCTGTCAACTAGAA 1508
Qy   1141 TTGTATGAGAAGCCCTACAGGATTTATGAAGCGGCACTTAAGATTGATTCATCAACAAA 1200
Db   1509 TTGTATGAGAAGCCCTACAGGATTTATGAAGCGGCACTTAAGATTGATTCATCAACAAA 1568
Qy   1201 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAATTTCAAGGAAACAGAACTAAATCT 1260
Db   1569 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAATTTCAAGGAAACAGAACTAAATCT 1628
Qy   1261 TAA 1263
Db   1629 TAA 1631

RESULT 2
AX833451 LOCUS AX833451 1559 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 575 from Patent EP1347046.
ACCESSION AX833451
VERSION AX833451.1 GI:39919586
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
Masuko, Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 575 24-SEP-2003;
Research Association for Biotechnology (JP)
FEATURES
source 1..1559
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      82.9%; Score 1047.4; DB 6; Length 1559;
Best Local Similarity 99.9%; Pred. No. 5e-195;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1  ATGCCTCTTCAGGTAGGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTCTGTCT 60
Db   43  ATGCCTCTTCAGGTAGGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTCTGTCT 102

Qy   61  CTGCCCTCAAAGCGGTGTGCGTCAGAGACACGACGACGTTCCTGCACGGAAAACTATCTG 120
Db   103 CTGCCCTCAAAGCGGTGTGCGTCAGAGACACGACGACGTTCCTGCACGGAAAACTATCTG 162
```

Qy	121	AAGGTCACATTTCTCCATTTTATTTGAGGCAATTTCTTTATGCTCCCATAGACGATGAG	180
Db	163	AAGGTCACATTTCTCCATTTTATTTGAGGCAATTTCTTTATGCTCCCATAGACGATGAG	222
Qy	181	AGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCCACCTTGATATAAAAAAGAACGG	240
Db	223	AGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCCACCTTGATATAAAAAAGAACGG	282
Qy	241	GCCATGTGGGAGACCCCTTTCTGTGTGACGGGTGTGTGACAAAGAGATGATGCAAGAAATTAGA	300
Db	283	GCCATGTGGGAGACCCCTTTCTGTGTGACGGGTGTGTGACAAAGAGATGATGCAAGAAATTAGA	342
Qy	301	GAATAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTTACAGAGCAAAAGCTCTCA	360
Db	343	GAATAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTTACAGAGCAAAAGCTCTCA	402
Qy	361	GCAAGCGGGAAGATCAAAAATACCACTTAAGTGTGATGATGAAGATTTGAAGAGAGAG	420
Db	403	GCAAGCGGGAAGATCAAAAATACCACTTAAGTGTGATGATGAAGATTTGAAGAGAGAG	462
Qy	421	AGGAAAAAATAGAGATATGAAGAAAAATGAACGGATAAAAGCCACTTAAGCAATTGGAA	480
Db	463	AGGAAAAAATAGAGATATGAAGAAAAATGAACGGATAAAAGCCACTTAAGCAATTGGAA	522
Qy	481	GCCTGGAAAGATATCAAGAAAGAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA	540
Db	523	GCCTGGAAAGATATCAAGAAAGAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA	582
Qy	541	TTATGTCAAAAAGAAAAAGCAATTAAGAAAGCAAAAAAATAAATAATAAGATCTT	600
Db	583	TTATGTCAAAAAGAAAAAGCAATTAAGAAAGCAAAAAAATAAATAATAAGATCTT	642
Qy	601	ACTGAAATTTGGCATCTAGAAATCTTCTCTCAAAAGGGAGAAAATTCAGAAAAATATATTT	660
Db	643	ACTGAAATTTGGCATCTAGAAATCTTCTCTCAAAAGGGAGAAAATTCAGAAAAATATATTT	702
Qy	661	ACTGAAAGTTAAAGGAAGACAGTATCTCTGCTCTGCTCTGCTGTCAGATTTAAATC	720
Db	703	ACTGAAAGTTAAAGGAAGACAGTATCTCTGCTCTGCTCTGTCAGATTTAAATC	762
Qy	721	AACCTTTACCCCTCGAGTATTCACACAGCTCTTCTGTGAATCACAAGTACAGAGAGAGAG	780
Db	763	AACCTTTACCCCTCGAGTATTCACACAGCTCTTCTGTGAATCACAAGTACAGAGAGAGAG	822
Qy	781	GAGTGGCTACACAAACAGCTGAGGACGACGAAGCAATGAATCTGACATAGCTGAATCTT	840
Db	823	GAGTGGCTACACAAACAGCTGAGGACGACGAAGCAATGAATCTGACATAGCTGAATCTT	882
Qy	841	TGCGATTTAAAGAAAGAAAGAACCCAGATGGTTGAAAGATTAAGGGAACCAATTTG	900
Db	883	TGCGATTTAAAGAAAGAAAGAACCCAGATGGTTGAAAGATTAAGGGAACCAATTTG	942
Qy	901	TTTTCACGGAAGAACTATTTGCGAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT	960
Db	943	TTTTCACGGAAGAACTATTTGCGAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT	1002
Qy	961	AATAAGATGCCACTATTTGTTGAACCGGCTGCTTCCACCTCAAAACTAAAAAATTTA	1020
Db	1003	AATAAGATGCCACTATTTGTTGAACCGGCTGCTTCCACCTCAAAACTAAAAAATTTA	1062
Qy	1021	CACAAGGCTATGAAGATCTTCTTAAGGC	1049
Db	1063	CACAAGGCTATGAAGATCTTCTTAAGGC	1091
RESULT 3	AK095201		
LOCUS	AK095201	1559 bp	mRNA linear PRI 30-JAN-2004
DEFINITION	Homo sapiens	CDNA FLJ37882	fis, clone BRSTN2000536.
ACCESSION	AK095201		
VERSION	AK095201.1	GI:21754405	
KEYWORDS	oligo capping; fis	(full insert sequence).	
SOURCE	Homo sapiens	(human)	

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayaashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Kamiyama, K., Katsuya, T., Kikawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuya, T., Sato, K., Tanikawa, M., Yanazaki, M., Nishimura, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoaka, S., Chiba, S., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Toqiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musahino, K., Yuuki, H., Oshima, A., Saaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Koniyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Nakajima, Y., Mizuno, T., Shinaga, M., Sasaki, M., Togaashi, T., Nagase, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizuehima-Sugano, J., Sato, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nagase, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

JOURNAL

Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Watanabe, M., Murakami, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

Unpublished
NEDO human cDNA sequencing project
3 (bases 1 to 1559)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

Location/Qualifiers

1. 1559

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRSTN2000536"

/tissue type="subthalamic nucleus"

/clone lib="BRSTN2"

/note="cloning vector: pME18SFL3"

43. .1173

/note="unnamed protein product"

/codon_start=1

FEATURES

source

CDS

/protein_id="BAC04498.1"
/db_xref="GI:21754406"
/translation="MPQVSDYSWQTKTAVFLSLPLKGVGVDRDVFCTENYLKNF
PPLFEALYAPIDDESSKAGNDITVFLYKKEAMWETLSVTGVDKEMMORIREK
SILQOERAEATEAKAARBDQKVALSVMMKI EBERKKI EDKEMNERIKATKALE
AWKEYORKEAQKIOREELCKQEKQIKYKSLRNLAENLAPKGRNSEN
IFTEKLKEDSIPAPRSVSGTKINFTPRVPTALRESOVAEEBEHLHQAARAMND
IAELCDLKEBEKPEWLKDKGNKLFPATENYLAALINAYNLAKMLPILYNRAACH
LKLNHLHKAIEDSSKAYRIMRHLRLHPTKLYKJMLDRFGM"

ORIGIN

Query Match 82.9%; Score 1047.4; DB 9; Length 1559;
Best Local Similarity 99.9%; Pred. No. 5e-195;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGCCTCTTCAGCTTACGATTACAGCTGGCAGACGAGACTGCGGCTTTCTCTCT	60
Db	43	ATGCCTCTTCAGCTTACGATTACAGCTGGCAGACGAGACTGCGGCTTTCTCTCT	102
Qy	61	CTGCCCTCAAAGCGTGTGGTCAGACACGACGCTGTTCGACGGAAACTATCTG	120
Db	103	CTGCCCTCAAAGCGTGTGGTCAGACACGACGCTGTTCGACGGAAACTATCTG	162
Qy	121	AAGTCAACTTCTCCTCAATTTTATTTGAGGCATTTCTTTATGCTCCCATAGCAGT	180
Db	163	AAGTCAACTTCTCCTCAATTTTATTTGAGGCATTTCTTTATGCTCCCATAGCAGT	222
Qy	181	AGCAGCAAGCAAGATTGGGAATGACACCATTTGTTACCTTTGTATATAAAAGAGCG	240
Db	223	AGCAGCAAGCAAGATTGGGAATGACACCATTTGTTACCTTTGTATATAAAAGAGCG	282
Qy	241	GCCATGTGGAGACCTTTCTGTACCGGTGTTGACAAAGAGATGATGCAAGATTTAGA	300
Db	283	GCCATGTGGAGACCTTTCTGTACCGGTGTTGACAAAGAGATGATGCAAGATTTAGA	342
Qy	301	GAAATATCTATTTTACAAGCACAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCA	360
Db	343	GAAATATCTATTTTACAAGCACAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCA	402
Qy	361	GCAAGCGGGAAGATCAAAATACGCCTAAGTGTCTATGATGAAGATTGAAGAGAGAG	420
Db	403	GCAAGCGGGAAGATCAAAATACGCCTAAGTGTCTATGATGAAGATTGAAGAGAGAG	462
Qy	421	AGGAAAAATAGAGATATGAAGAAATGAACGGATTAAGCCACTTAAGCATTGGAA	480
Db	463	AGGAAAAATAGAGATATGAAGAAATGAACGGATTAAGCCACTTAAGCATTGGAA	522
Qy	481	GCCTGGAAAGAAATATCAAGAAAAAGCTGAGGACAAAAAATTTACAGAGAGAGAAA	540
Db	523	GCCTGGAAAGAAATATCAAGAAAAAGCTGAGGACAAAAAATTTACAGAGAGAGAAA	582
Qy	541	TTATGTCAAAAGAAAAAGCAATTTAAAGAGAGAAAAAATATAAATATAAGAGCTTT	600
Db	583	TTATGTCAAAAGAAAAAGCAATTTAAAGAGAGAAAAAATATAAATATAAGAGCTTT	642
Qy	601	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAATATATTT	660
Db	643	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAATATATTT	702
Qy	661	ACTAGAAATTTAAAGGAAGACAGTATTTCTGCTCTCTGTTGGCAGTATTAATAATC	720
Db	703	ACTAGAAATTTAAAGGAAGACAGTATTTCTGCTCTCTGTTGGCAGTATTAATAATC	762
Qy	721	AACCTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGAG	780
Db	763	AACCTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGAG	822
Qy	781	GAGTGGCTACACAAACAGCTGAGCAGAGAGCAATGAATCTGACATAGCTGAACCTT	840
Db	823	GAGTGGCTACACAAACAGCTGAGCAGAGAGCAATGAATCTGACATAGCTGAACCTT	882
Qy	841	TGCCATTTAAAGAGAGAGAGAGAACCCAGAAATGGTTGAAGGATAAAGGAGAGAAATTTG	900

Db	883	TGCGATTTAAAGACAGAGAAAAAGAACCCAGAAATGGTTGAAGATAAAGGAAACAATTTG	942
Qy	901	TTTGCAACGAAAACTATTGGCAGCTATCAATGCATATATTTAGCCATTAAGACTAAAT	960
Db	943	TTTGCAACGAAAACTATTGGCAGCTATCAATGCATATATTTAGCCATTAAGACTAAAT	1002
Qy	961	AATAAGATGCACTATTGTATTGAAACGGGCTGCTTGCCACCTAAAAAACTTA	1020
Db	1003	AATAAGATGCACTATTGTATTGAAACGGGCTGCTTGCCACCTAAAAAACTTA	1062
Qy	1021	CACAAGCTATTGAAGATTTCTTAAGGC	1049
Db	1063	CACAAGCTATTGAAGATTTCTTAAGGC	1091

RESULT 4	BC062564	1468 bp	mRNA	linear	PRI 28-NOV-2003
LOCUS	clone MGC:70618 IMAGE:5163101				1, mRNA (cdna)
DEFINITION	Homo sapiens dyslexia susceptibility 1 candidate 1, complete cds.				
ACCESSION	BC062564				
VERSION	BC062564.1	GI:38565951			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1468)				
	Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Schuler, G.D., Krausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sapleton, T.E., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A., 2002				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1468)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: gcgbs-x@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@hgrl.nih.gov Ahter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,				

human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)
 12477932
 2 (bases 1 to 1946)
 REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Grimwood, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 44 Row: b Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

source
 1..1946
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:31465 IMAGE:4483015"
 /tissue_type="Mammary tumor. Metallothionien-TGF alpha
 model. 10 month old virgin mouse. Taken by biopsy."
 /clone_lib="NCI CGAP Mam1"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 gene
 1..1946
 /gene="Dyxlcl1"
 /note="synonym: EKN1"
 /db_xref="LocustID:67685"
 /db_xref="MGI:1914935"
 CDS
 48..1310
 /gene="Dyxlcl1"
 /codon_start=1
 /product="dyslexia susceptibility 1 candidate 1 homolog"
 /protein_id="AAH26462.1"
 /db_xref="GI:20071434"
 /db_xref="LocustID:67685"
 /db_xref="MGI:1914935"
 /translation="MPYRVSEFSWQOTPATIFSLPLRGVCRDADVFCGESYLKVPF
 PPFLFELFAPIDGSKAKIGNDTILFTLYKKEPVLNDSLNPQVDKEMQRIERK
 SILQAKAEATEAKAVAKREDQRYALGEMMKIEBERKKIEMDKENRKKATSELE
 AWKECKQADQKQVORKEPPEKQAEETKALPKGLPRKAPPTPLTRGNWENIF
 PEKLKEDRPAPRAGSIQISFTPRVPTALRESQVAEEBWLHKQAEARRAMSTDLP
 EFPDLKEERNPDWLKGNKLPATENYLAADVAYNLAIRLNCKIPLLYLNRAACHLK
 LKHLKALIESSKALELLTPPVADNANRMAKHVRGTAFCQLELYVEGLQDYEAALK
 IDPANTVVQNDAEKIRNLIQGTALKSRD"

ORIGIN

Query Match 67.0%; Score 846; DB 10; Length 1946;
 Best Local Similarity 80.1%; Pred. No. 1.3e-155;
 Matches 1009; Conservative 0; Mismatches 245; Indels 6; Gaps 1;
 Oy 1 ATGCTCTTACAGTTAGCGATTACAGCTGGCAGCAGCAGAACTGCGGTCTTCTGTCT 60
 Db 48 ATGCGAGTGCAGTGAGGAATTCAGCTGGCAGCAGCAGCGCGCCACGATCTCTGTGCG 107
 Oy 61 CTGCCCCCTCAAGGGCTGTGCGTCAGACACGCGACGCTGTTCTGCACGGAAACTATCTG 120

Db 108 CTGCTCTGCGGGCGTCTGCGTGGCGATGCTGACGTATTCCTGTGGGAAAGTTACTCTG 167
 Oy 121 AAGGTCAACTTCTCTCCATTTTATTGAGGCATTTCTTATGCTCCCATAGACGATGAG 180
 Db 168 AAGGTAACTTCTCTCCATTTTATTGAGCTGTTCTCTATGCTCCCATAGATGATGGG 227
 Oy 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTTCACCTTTGTATAAAAAAGAGCG 240
 Db 228 AAGAGCAACCGAAGATTGGAAATGACACGATTTCTTTTCACTTTGTATAAAGGAGCCCA 287
 Oy 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCCAAAGATTAGA 300
 Db 288 GTTCTGTGGATAGCCCTTTCTATGCGGGTGTGATAAAGAGATGATGCAGAGAAATAGA 347
 Oy 301 GAAAAATCTATTTTACAGCACAAGAGAGAGCAAAAGAAAGAACTACAGAGCAAAAGCTGCA 360
 Db 348 GAAAAATCTATCTTGGCAAGCAGGCAAAAGAAAGAGGCGCCACAGAAAGCAAAAGCTGTT 407
 Oy 361 GCAAGCGGGAAGATCAAAATACGCACCTAAGTGTCTATGATGAAGATTGAAGAAGAAGAG 420
 Db 408 GCCAAGCGAAGACCCAGATAGCCACTAGGCGAGATGATGAAGATTGAAGAAGAAGAG 467
 Oy 421 AGGAAAAAATAGAGATATGAAAGAAATGAAACGGATATAAAGCCACTAAAGCATTTGGAA 480
 Db 468 AGGAAAAAATAGAGATATGAAAGAAATGAAACGGAAAGAAAGCACTAGCGAATTAGAA 527
 Oy 481 GCCTGGAAGAAATATCAAGAAAGCTGAGAGCAAAAAAATTTAGAGAGAAGAGAAA 540
 Db 528 GCGTGAAGAAATGCTCAAAAGAAAGCTGACGGCAAAAAAGAGTCCAGAGAGAAAGAGAAA 587
 Oy 541 TTATGTCAAAAAGAAAGCAAAATTAAGAAGAGGAGAAAAAATAAATAATAAGAGTCTT 600
 Db 588 -----CCGCCCGAGAGAAAGCAAGCTGGAAGAGACCAAAGCTTAAACCTTCGGGGTTG 641
 Oy 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAAGGAGAGAAATTCAGAAAAATATATTT 660
 Db 642 CCCCAGAGGCCCCCAACCCACTCGCTCCCAAGAGAGAGGAGGAATTCGGGAAACATATTT 701
 Oy 661 ACTGAGAAGTTAAAGGAAGACGATTTCTGCTCTCGCTCTGTTGGCAGTATTAAATC 720
 Db 702 CCTGAGAAGTTAAAGGAAGACGAGATCCCTCGCGCTCGCTCCGCTGGCAGTATTCAAATC 761
 Oy 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTTCTGCTGAATCACAAAGTAGCAGAGAGAG 780
 Db 762 AGCTTTACCCCTCGAGTGTTCCTCAACAGACTTCGGGAATCCCAAGTTCGCAAGAGAGAG 821
 Oy 781 GAGTGGCTACACAAAGCTGAGGCCACGAAGCAATGAATACTGACATAGCTGAATCTT 840
 Db 822 GAGTGGCTGCATATAAAGCAGAGACGAGAGCCATGAGCACTGACCTTCTCTGAATTC 881
 Oy 841 TGCATTTAAAGAGAGAGAAAGAACCCAGAAATGGTTGAAGATTAAGGAAACAAATTTG 900
 Db 882 TTTGACTTAAAGAGAGAGAGAGGAATCCAGACTGGTGTGAAAGAACAAGGGGAAACAAATTTG 941
 Oy 901 TTTGCAAGGAAACTATTTGGCAGCTATCAATCATATAATTTAGCCATAGAACTAAAT 960
 Db 942 TTTGCAACAGAAAACTATTTGGCAGCGGTGATGCATATATAATTTAGCCATAGACTGAAC 1001
 Oy 961 AATAAGATGCCACTATTTGTATTTCGAACGGGCTCTTGGCCACCTTAAATAAAGAACTTA 1020
 Db 1002 TGTAAGATCCCATATTTGTATTTCGAATCGGGCTCTTGGCCACCTCAAAATTAAGAAACCTTA 1061
 Oy 1021 CACAAGGCTATGAAAGATTTCTTAAGGCACTGGAATTTATGATGCCACCTGTTTACAGAC 1080
 Db 1062 CACAAGGCCATCGAGGACTCTTCTTAAGGCACTAGAGTTATTTGACACCACCTGTTGCTGAC 1121
 Oy 1081 AATGCTTAATCAAGAATGAAGGACATGTACGAGTGGAAAGAGAGATTTCTGTCAACTAGAA 1140
 Db 1122 AATGCCAATGCAAGATGAAGGACACGCTCCGAGAGGACAGGCTTCTGTCAACTAGAA 1181
 Oy 1141 TTGTATGTAGAAAGGCTTACAGGATTATGAAGCGGCACCTTAAGATTGATTCCTCAACAAA 1200

Db 460 AGAAACGATTGTAAGATGTCACAGAGGAGGAAAGAAAGCAACAGAGGAACTGGAG 519
Qy 481 GCCTGGAAAGATATCAAGAAAGCTGAGAGGACAAAGAAATTCAGAGAGAGAGAAA 540
Db 520 AAGTGGAAAGAAACAGAGGATTCAGAAGTTAAATGAAATGATCCAAAAAACA---- 575
Qy 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAAAGAAATTAAGAGTCTT 600
Db 576 -----AATAGCAAGAGATGAAGAAATAAGCAATACACAGGGTCTTGAAC 628
Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTGGCTCCAAAGGGAGAAATTCAGAAATATATTT 660
Db 629 ACTGCTCA-----ACTGCAACGACTGTACAAATGACCGTACAAAT 669
Qy 661 ACTGAGAAGTTAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTGGCAGTATTAATATC 720
Db 670 AACGTCGAAGCAAGAAAGAAATTTGCTCTCTACTCGGTCCACTGGGTGCAATTCAGATC 729
Qy 721 AACTTTTACCCCTCCAGTATTCCTCAACAGCTCTTCGTGAATCACAAGTAGCAGAGAGGAG 780
Db 730 CAGTTTACCTCGTGCTCTTCTTACAGCACTCAGAGATCCCGTGTGGCAGAGAGAGAA 789
Qy 781 GAGTGGCTACACAAACAGCTGAGGCACGAGCAATGAATACTGCATATAGCTGAACCTT 840
Db 790 GAATGGTTACACAGCAGGCGGAGCCGAGAGCCCACTAATACAGTTGATCTTGGACTT 849
Qy 841 TGCATTTAAGAGAGAGAAAGAACCCAGAAATGGTTGAAGATAAGGAACAAATTTG 900
Db 850 GAAGACTTGAACCGAAGAGGAGAAAGAAATCCAGAGTGGCTTAAAGATAAAGGAAACAGCTC 909
Qy 901 TTTCCAAACGGAAGAACTATTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960
Db 910 TTTCCAGCTGAGACTATCTTGGCTGTGCTGTGAATGCTCTACACCTGGCAATACGGTTAAT 969
Qy 961 AATAAGATGCCACTATTTGTAATTTGAACCGGCTCTTGGCCACCTTAAAGAACTTAA 1020
Db 970 AATAAGATTCAGGACTGTATCTAAATCGATCTCGCTGTCACTTTAAACTAAGAAACCTT 1029
Qy 1021 CACAAGCTATTGAAGATCTTCTAAGGCACCTGGAATTTATGATGCCACCTGTTACAGAC 1080
Db 1030 CATAAAACAATTTGAGGACTCATCAAGGCTTTTGGAGCTACTTACACCACCGGTTCCAGGC 1089
Qy 1081 AATCCTAATCAAGAAATGAAGGCACATGTACGAGTGGAGCAGCACTTGTCAACTAGAA 1140
Db 1090 AATCCAGTCCAGAAATTAAGGCTCATGTAGGCGGGGACAGCTTCTGTGAATTTGAA 1149
Qy 1141 TTGTATGTAGAGCCCTACAGGATTTAAGACCGGCACCTTAAGATTGATCCATCCAAACAA 1200
Db 1150 CTTTATGTAGAAGGCTCCAAAGACTACGAGCTGCCCTTAAATTTGATCCCAACCAATCAA 1209
Qy 1201 ATTGTACAAATTTGATGCTGAGAGAAATTCGGNATGTAATCA 1242
Db 1210 AACATGAATCTGATGCTGAAAGAAATTCGACGGGTCAATCCAA 1251

RESULT 7
AR413597 LOCUS 489 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1234 from patent US 6639063.
ACCESSION AR413597
VERSION AR413597.1 GI:40168707
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 489)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 1234 28-OCT-2003;
FEATURES Location/Qualifiers
source 1. .489
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 34.3%; Score 432.6; DB 6; Length 489;
Best Local Similarity 99.8%; Pred. No. 1.3e-74;
Matches 432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAAGACTGGCGTCTTCTGTCT 60
Db 57 ATGCTCTTTCAGGTTAGCGATTACAKCTGGCAGCAGACGAAAGACTGGCGTCTTCTGTCT 116
Qy 61 CTGCCCCCAAAGCGGTGTCGTGAGACACGCGACGTGTCGACGGAAGAACTATCTG 120
Db 117 CTGCCCCCAAAGCGGTGTCGTGAGACACGCGACGTGTCGACGGAAGAACTATCTG 176
Qy 121 AAGTCAACTTTCCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGAGATGAG 180
Db 177 AAGTCAACTTTCCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGAGATGAG 236
Qy 181 AGCAGCAAAACCAAGATTGGGAATGACACCATTTGTCTTCCACCTTGTATATAAAAGAAAGCG 240
Db 237 AGCAGCAAAACCAAGATTGGGAATGACACCATTTGTCTTCCACCTTGTATATAAAAGAAAGCG 296
Qy 241 GCCATGTGGAGACCCCTTTCTGTGACCGGTGTTGACAAAGAGATGATGCAAAAGAAATAGA 300
Db 297 GCCATGTGGAGACCCCTTTCTGTGACCGGTGTTGACAAAGAGATGATGCAAAAGAAATAGA 356
Qy 301 GAAAAATCTATTTTACAGCAGCAGAGAGAGCAAAAGAAAGCTACAGAGCAAAAGCTGCA 360
Db 357 GAAAAATCTATTTTACAGCAGCAGAGAGAGCAAAAGAAAGCTACAGAGCAAAAGCTGCA 416
Qy 361 GCAAAGCGGAAGATCAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAGAG 420
Db 417 GCAAAGCGGAAGATCAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAGAG 476
Qy 421 AGGAAAAAATAG 433
Db 477 AGGAAAAAATAG 489

RESULT 8
AX970431 LOCUS 489 bp DNA linear PAT 15-JAN-2004
DEFINITION Sequence 1234 from Patent EP1104808.
ACCESSION AX970431
VERSION AX970431.1 GI:40977781
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE ESTs and encoded human proteins
JOURNAL Patent: EP 1104808-A 1234 06-JUN-2001;
Genset (FR)
FEATURES Location/Qualifiers
source 1. .489
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
57. .>488
/note="unnamed protein product"
/codon_start=1
/protein_id="CAFI5539.1"
/db_xref="GI:40977782"
/translation="MPLQVSDYXWQQTAVFLSLPLKGVCVDRDVFCTENYLKVP
PPFLFEAFLYAPIDDESSKAKIGNDIIVFTLYKKAAMWETLSVTGDKEMQRIK
SILQAQERAKEATEAKAAKREDQKVALSVMMKIEBERKKI"

ORIGIN
Query Match 34.3%; Score 432.6; DB 6; Length 489;
Best Local Similarity 99.8%; Pred. No. 1.3e-74;
Matches 432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;


```
Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAAGACTGCGGTCTTCTGTCT 60
Db 57 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAAGACTGCGGTCTTCTGTCT 116
Qy 61 CTGCCCCCTCAAGGGCGTGTGGTTCAGACACGACGCGTGTCTGACCGGAAACTATCTG 120
Db 117 CTGCCCCCTCAAGGGCGTGTGGTTCAGACACGACGCGTGTCTGACCGGAAACTATCTG 176
Qy 121 AAGGTCAACTTTCCTCCATTTTATTTGAGGCAATTTCTTTATGTCTCCCATAGACGATGAG 180
Db 177 AAGGTCAACTTTCCTCCATTTTATTTGAGGCAATTTCTTTATGTCTCCCATAGACGATGAG 236
Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCAATTTCTTACCTTGTATATAAAGAAAGCG 240
Db 237 AGCAGCAAGCAAGATTGGGAATGACACCAATTTCTTACCTTGTATATAAAGAAAGCG 296
Qy 241 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAAATTAGA 300
Db 297 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAAATTAGA 356
Qy 301 GAAAAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db 357 GAAAAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 416
Qy 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 420
Db 417 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 476
Qy 421 AGGAAAAAATAG 433
Db 477 AGGAAAAAATAG 489
```

```
RESULT 9
LOCUS BD109150 489 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD109150
VERSION BD109150.1 GI:23203968
KEYWORDS JP 2002010789-A/1227.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 489)
EDWARDS, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 1227 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/1227
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS 57..488.
FEATURES
source
1..489
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 34.3%; Score 432.6; DB 6; Length 489;
Best Local Similarity 99.8%; Pred. No. 1.3e-74;
```

```
Matches 432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAAGACTGCGGTCTTCTGTCT 60
Db 57 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAAGACTGCGGTCTTCTGTCT 116
Qy 61 CTGCCCCCTCAAGGGCGTGTGGTTCAGACACGACGCGTGTCTGACCGGAAACTATCTG 120
Db 117 CTGCCCCCTCAAGGGCGTGTGGTTCAGACACGACGCGTGTCTGACCGGAAACTATCTG 176
Qy 121 AAGGTCAACTTTCCTCCATTTTATTTGAGGCAATTTCTTTATGTCTCCCATAGACGATGAG 180
Db 177 AAGGTCAACTTTCCTCCATTTTATTTGAGGCAATTTCTTTATGTCTCCCATAGACGATGAG 236
Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCAATTTCTTACCTTGTATATAAAGAAAGCG 240
Db 237 AGCAGCAAGCAAGATTGGGAATGACACCAATTTCTTACCTTGTATATAAAGAAAGCG 296
Qy 241 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAAATTAGA 300
Db 297 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAAATTAGA 356
Qy 301 GAAAAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db 357 GAAAAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 416
Qy 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 420
Db 417 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 476
Qy 421 AGGAAAAAATAG 433
Db 477 AGGAAAAAATAG 489

RESULT 10
LOCUS CQ685075 464 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 30001 from Patent WO2070737.
ACCESSION CQ685075
VERSION CQ685075.1 GI:42211607
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Liew, C.C., Marshall, W.E. and Zhang, H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 30001 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source
1..464
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 31.8%; Score 401.6; DB 6; Length 464;
Best Local Similarity 98.3%; Pred. No. 1.5e-68;
Matches 404; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAAGACTGCGGTCTTCTGTCT 60
Db 49 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAAGACTGCGGTCTTCTGTCT 108
Qy 61 CTGCCCCCTCAAGGGCGTGTGGTTCAGACACGACGCGTGTCTGACCGGAAACTATCTG 120
Db 109 CTGCCCCCTCAAGGGCGTGTGGTTCAGACACGACGCGTGTCTGACCGGAAACTATCTG 168
Qy 121 AAGGTCAACTTTCCTCCATTTTATTTGAGGCAATTTCTTTATGTCTCCCATAGACGATGAG 180
Db 169 AAGGTCAACTTTCCTCCATTTTATTTGAGGCAATTTCTTTATGTCTCCCATAGACGATGAG 228
```

181 AGCAGCAAGCAAGATTGGGAATGACACCATCTCTTACCTTGTTATATAAAGAGAGCG 240
 |||||
 Db 229 AGCAGCAAGCAAGATTGGGAATGACACCATCTCTTACCTTGTTATATAAAGAGAGCG 288
 |||||
 Qy 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTACAAAGAGATGATGCCAAAGATTAGA 300
 |||||
 Db 289 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTACAAAGAGATGATGCCAAAGATTANA 348
 |||||
 Qy 301 GAAATAATCTATTATCAAGCAAGAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 360
 |||||
 Db 349 GAAATAATCTATTATCAAGCAAGAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 408
 |||||
 Qy 361 GCAAGCGGAGAGATCAAAATACGCCTAAGTCTCATGATGAGATTGAA 411
 |||||
 Db 409 GCAAGCGGAGAGATCAAAATACGCCTAAGTCTCATGATGAGAGAGAA 459
 |||||

RESULT 11
 BC065881
 LOCUS
 DEFINITION Danio rerio zgc:77853, mRNA (cDNA clone MGC:77853 IMAGE:7002321), complete cds.
 ACCESSION BC065881
 VERSION BC065881.1 GI:41389019
 KEYWORDS MGC.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skaleka, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

TITLE
 JOURNAL 2 (bases 1 to 1443)
 PUBMED Strausberg, R.
 REFERENCE Direct Submission
 AUTHORS Submitted (26-JAN-2004) National Institutes of Health, Mammalian
 TITLE Gene Collection (MGC), Cancer Genomics Office, National Cancer
 JOURNAL Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK Contact: MGC help desk
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: Len Zon, Harvard
 cDNA Library Preparation: Open Biosystems
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,

Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabbu, Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skaleka, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 147 Row: p Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES
 source
 1. .1443
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:77853 IMAGE:7002321"
 /tissue_type="Whole, adult zebrafish"
 /clone_lib="NIH_ZGC_10"
 /lab_host="DH10B"
 /notes="Vector: pExpress1"
 1. .1443
 /genes="zgc:77853"
 /notes="synonym: MGC77853"
 /db_xref="LocusID:402988"
 19. .1281
 /genes="zgc:77853"
 /codon_start=1
 /product="hypothetical protein zgc:77853"
 /protein_id="AAH65881.1"
 /db_xref="GI:41389020"
 /db_xref="LocusID:402988"
 /translation="MPLIVRDHTVNLSTVIVPLKAVKTANVHICTDDYLKVSFPPFLFVFLFASINEEKSEARINGAAVFLQKRDEMWEQLCTNIDKEKQKQIREQA LKQVQKAVKSKTKATIQOEKKYALETMKLESEERERIQKRKEVCAATAE LDF WRTOKTKAGNENQIKQRTSHDNKPVQPKIKHANTS PADTALVTDNTITGOMRS KOTKDLPA PRSAGCIQISFTPRVPTALRESRVPEEBEWMKQAEARRAIDTDLASL DDILTEBERNPDMLKDKGDLFMAGNPLAAVNAYNLATKLNRKMPALFNSRAACHLKLRLNHLKEDSSQALELITPAVSANASARKLKAHVARGTAFQCLQLYVEGLQDYQKALEND PHNALRADTEQIRELIQGSTPRSKDEN"
 ORIGIN
 Query Match 30.9%; Score 390.2; DB 5; Length 1443;
 Best Local Similarity 58.0%; Pred No. 2.4e-66;
 Matches 733; Conservative 0; Mismatches 518; Indels 12; Gaps 2;
 Qy 1 ATGCTCTTTCAGGTTACGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTCTGTCT 60
 Db 19 ATGCCGCTGATAGTAAGAGATCACATCGGACACAAACCTGAGCACAGTTTACATCAGC 78
 Qy 61 CTGCCCTCAAGCGTGTGGTTCAGACACGAGACGCTGTTCTGCAGGAAACCTATCTG 120
 Db 79 GTCCCTTTAAAGCAGTGAAACCGCAAAATGTACACATCATCTCGACGGACGACTACCTG 138
 Qy 121 AAGGTCAACTTTCCTCCATTTTATTTTGGGCACTTCTTTATGTCTCCATAGACGATGAG 180
 Db 139 AAGTCTAGTTTCCACCCTTCTCTTCAGGTTTCTTATTCGGCTCAATAAATGAGGAG 198
 Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCATCTGTCTTCACCTTTGTATAAAGAGAGCG 240
 Db 199 AAAAGTAGCCAGGATTGGAAACGGTGCTGCAGTTTCTACTCTTCAGNAGGAGAGAT 258
 Qy 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCCAAAGATTAGA 300
 Db 259 GAAATGTGGGAGCAGCTCTGT---ACAAACATCGATAAAGAGAAACAGAACCAATTCGA 315
 Qy 301 GAAATAATCTATTATCAAGCAAGAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 360
 Db 316 GAACAGCCGTTCTCAAAGTTTCAGAGAAAGCAGTGGNAAAGTCTAAACCCAGGCCACC 375
 Qy 361 GCAAAGCGGAGAGATCAAAATACGCCTAAGTCTCATGATGAGATTGAAAGAGAGAGAG 420
 |||||

376 CGGATTTCAAGGAGAAAGTATCGCTGGAAACCATGATGAAGCTTGAGCGGAGGAG 435
421 AGGAAAAAATAGAAAGATATGAAGAAATGAACGATAAAGCCACATTAAGCATTTGAA 480
436 CGGAGAGAGATTCAGAGAGGAAGATGAAGTGTGTGCGAGAGCCACACGACGAGCTGGAT 495
481 GCCTGGAAGATATCAAGAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 540
496 TTCTGGAGAGAAACACAGAGAAAAACGCCAAGGGAAACGAGAACAGATAAAGAGCAG 555
541 TTATGTCAAAAGAAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
556 AGAACCCAGCACATGACACAAACCGGTCTCTAGAGAAATTAAGCATGCAACACTTCA 615
601 ACTAGAAATTTGGCATCTAGAAATCTTGTCCAAAGGGAGAGAAATTCAGAAAAATATTT 660
616 CTTGAGAGACAGCAT-----TGGTGAACCCGAGCAACACTATCACTGGTCAGATG 666
661 ACTGGAAGTTAAGGAGACAGTATTTCTGCTCTCTGCTCTGTTGGCAGTATTAATATC 720
667 AGGAGTAACAGAGAAACAAAGAGATCTCCCGCTCCAGATCTGCTGGCTGATTTCAGATC 726
721 AACTTTACCTTCAGATTTCCACAGCTCTTCTGTGAATCACAAGTAGCAGAGAGGAG 780
727 AGCTTCACTCAGAGATGTTTCCACCGGCTCAGAGAGTCTCGGCTCCCGAGGAGGAG 786
781 GAGTGGCTACAAACAAAGCTGAGCACAGAGAGCAATGAATCTGACATAGCTGAACCTT 840
787 GAGTGGTTGAAGAGCAGCGGAGCCAGGAGCAATAGACACAGATCTGCGAGAGCTG 846
841 TCGCATTTAAAGAGAGAGAAAGAACCCAGAGATGGTTGAAGGATAAAGGAAACAAATTTG 900
847 GACGACCTGACAGAGGAGAGAGAAACCCAGACTGGCTGAAAGACAAAGGAGATAAATTTG 906
901 TTGCAACGGGAAACTATTTGCGAGCTCAATGATATATTTAGCCATTAAGACTTAAT 960
907 TTTATGGCAGGAACTTTCTGCTGCTCAAGCTTATATCTGCGCATTAAGACTCAAC 966
961 AATAAGATGCCATATTGTTGAACCGGGCTGTTGCCACCTTAAACTAAAAAACTTA 1020
967 AGGAAGATGCGGGCTTTATTTCTAATCGAGCGGCTGTGCACTCAAACTGAGAAATCTC 1026
1021 CACAGGCTATTGAAGATCTTTTAAGGCACTGGAATTTATGATGCCACCTGTTACAGAC 1080
1027 CACAGGCACTTGAAGGACAGCTCTCAGCGCTTGAAGTTGCTGACGCTCGCGTCTCTCG 1086
1081 AATGCTAATCAAGAAATGAAGCACTCTAGAGCTGGAAGAGCAATCTGTCAACTAGAA 1140
1087 AACGCTTCAGCTCGAGCTGAAGCTCATGTGAGCGGGGAGACGCTTCTGTGAGCTCCAG 1146
1141 TTGTATGTAGAGGCTTACAGATTTAAGAGCGGCACTTAAGATTTGATCCATCCAAACAA 1200
1147 CTCTATGTGAAGGCTTCAAGACTATCAGAAGGCTCTAGAAATGGACCCACACAAATGCA 1206
1201 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAAATTCAGAGAACAGAACTAAATCT 1260
1207 GCACTGCGAGCGGATACCGAGAGATCGGTGAGCTCATACAGGGATCCACTCCAGATCT 1266
1261 TAA 1263
1267 AAA 1269

RESULT 12
CR387522
LOCUS CR387522 761 bp mRNA linear VRT 05-APR-2004
DEFINITION Gallus gallus finished cdna, clone CHEST505n19.
ACCESSION CR387522
VERSION CR387522.1 GI:46240281
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 761)
Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Graham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
Direct Submission
Submitted (05-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickens@hms.unist.ac.uk
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cdna
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cdna collection,
from a library constructed by Elizabeth Bosch. cdna was prepared
from RNA extracted from hearts, normalised, and poly A-trimmed.
EcORI-NotI cut cdna was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
FEATURES
Location/Qualifiers
1..761
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST505n19"
/clone_lib="CSEQCHN60"
/dev_stage="adult"
ORIGIN
Query Match 23.6%; Score 298; DB 5; Length 761;
Best Local Similarity 76.8%; Pred. No. 2.8e-48;
Matches 364; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 790 CACAAACAGCTGAGGACGAGAGCAATGAATGCTGATTTGCTGAGCTGGAAGATTG 849
Db 2 CAAAAACAGGAGAGAGCTCGAAGGACAAATAAGTGTGATTTGCTGAGCTGGAAGATTG 61
QY 850 AAAGAAG 909
Db 62 AAAG 121
QY 910 GAAACATTTTGGCAGCTATCAATGCAATATATTTAGCATAAGACTAAATAAATAGATG 969
Db 122 GGAGACTATCTTGAGCTGTAATGCAATATTAACCTTGCACTCCGCTTAAACAAATAGCTT 181
QY 970 CCATTTGTTATTTGAACCGGCTGCTTGCACCTTAAACCTTAAACAACTTACACAGGCT 1029
Db 182 CCATCTGTTACTTGAATCGTGTGCTTGCCTTAAAGCTGAGAGAAATTTACACAAAGCC 241
QY 1030 ATTGAAGATTTCTTAAAGGCACTGGAATTTATTTGATGCCACCTGTTTACAGACAAATGCTAAT 1089
Db 242 ATTGAAGATTTCTTAAAGGCACTGGAATTTATTTGATGCCACCTGTTTACAGACAAATGCTAAT 301
QY 1090 GCAAGAGTGAAGGACATGTACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1149
Db 302 GCTCGAGTAAAGCATATGTGAGAGCTGGCAGAGCTTTTGTGAGCTGGAAGCTGATGCT 361
QY 1150 GAAGGCTTACAGGATTTATGAGCGGCACTTAAGATTTGATCCATCCAAACAAATTTAGTACAA 1209
Db 362 GAAGGCTTCCAGGATTTATGAAGAGAGCTTCAAGATTTGATCTTAAACAAATTTATAGAA 421
QY 1210 ATTGATGCTGAGAGAGATTCGGAATTTGATTTCAAGAGAGAGAGAGAGAGAGAGAGAG 1263
Db 422 AAAGATGAGAGAGAGATTCGAGACACATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 475
RESULT 13
CR389904
LOCUS CR389904 1133 bp mRNA linear VRT 20-APR-2004
DEFINITION Gallus gallus finished cdna, clone CHEST271n13.
ACCESSION CR389904

```

VERSION      CR389904.1  GI:46428549
KEYWORDS
SOURCE       Gallus gallus (chicken)
ORGANISM     Gallus gallus
REFERENCE    1 (bases 1 to 1133)
AUTHORS      Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
              Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
              Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
              Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
              Tickle,C. and Wilson,S.A.
TITLE        Direct Submission
JOURNAL      Submitted (19-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
COMMENT      CB10 ISA, UK. E-mail enquiries: chickest@ms.umist.ac.uk
              BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
              sequencing project.
              This sequence is from the
              BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
              from a library constructed by Elizabeth Bosch. cDNA was prepared
              from RNA extracted from ovary, normalised, and poly A-trimmed.
              ECORI-NotI cut cDNA was then ligated into the vector. Vector:
              pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
              coli DH10B.
FEATURES     Location/Qualifiers
             1..1133
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="Layer-"
                /db_xref="taxon:9031"
                /clone="CHSR271n13"
                /clone_lib="CSEQRN19"
                /dev_stage="adult"
ORIGIN
Query Match      21.8%; Score 275; DB 5; Length 1133;
Best Local Similarity 66.4%; Pred. No. 8.8e-44;
Matches 395; Conservative 0; Mismatches 200; Indels 0; Gaps 0;
Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGCAGAGACTGGCGTCTTCTCTCT 60
Db 73 ATGCGCTGTGCTGCGGAGTACAGCTGGCGGAGCGCGCGGTACTCTCTCG 132
Qy 61 CTGCCCTCAAAGCGGTGTGCTCAGACACGAGCGTGTCTGCAGCGAAACATCTCTG 120
Db 133 CTGCCCGTGGCGGCTGAGGGTCAACGCCGCCCAACATCTCTGCACGCGGTACCTG 192
Qy 121 AAGTCACACTTTCCTCCATTTTATTGAGGCATTTCTTTATGCTCCCATAGCGATGAG 180
Db 193 AAGGTAGCGTCCCTCCCTTTTATTCGAAGCTGTTCTGTAAGTCTCTTATTTGATGAGACA 252
Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCAATGTCTTCACTTTGTATATAAAGAGAGCG 240
Db 253 AATAGCAGCAAGATCGCAATGGAGTCTGTTCTTCACTCTGTATATAAAGAGAGCG 312
Qy 241 GCCATGTGGAGACCTTTCTGTGCGGGTGTGACAAAGAGATGATGCAAGAAATTAGA 300
Db 313 GCCATGTGGAGTCCCTGGCTGTGCAAAATGTAACAGGAGGAACTGCAACGCTCTAAGA 372
Qy 301 GAAATACTATTTTACAGACACAGAGACCAAGAGAGTACAGAGCAAAAGCTGCA 360
Db 373 GAGAAATGCTGTCTGAAAGCACACGAAAGCAAAAGAGAGAGAGCAAGCAAAAGTT 432
Qy 361 GCAAAGCGGGAAGATCAAAAATAGCGCACTAAGTGTGTCATGATGAAGATTGAAGAAGAGAG 420
Db 433 ACBAACAGCAAGACATAAGAGATGCTTTTGGAGGCTACTAATGAAGCTAGAGAGAGAGAA 492
Qy 421 AGGAAAAAATAGAGATATGAAGAAATGAACGGATAAAGCCATTAAGACGTTGGAA 480
Db 493 AGAAAAAGAAATTGAAGATCTGAAAGAACAGAGAGACAGAAAGTGCCTAAGGAGTTGGAG 552
Qy 481 GCCTGGAAGAAATATCAAGAAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAGAA 540

```

```

Db 553 CTATGCGAGAGTCAGCAAAACGATGTTGAGAAAACAAAGAGGCTACAAAAGAAAGGGGAA 612
Qy 541 TTATGTCAAAAGAAAGCAAAATTAAGAGGAGGAAGAAAAAATAAATATTAAGA 595
Db 613 CTACATGAAGAGTAGAGCAACTAAAGAGGAGAGAGAAAAAGAGAAAAAACAATA 667

```

```

RESULT 14
AC022083/c 114149 bp DNA linear PRI 12-MAY-2001
LOCUS Homo sapiens chromosome 15 clone CTD-2137J4 map 15q21.3, complete
DEFINITION
AC022083
AC022083.6 GI:114029048
VERSION Homo sapiens (human)
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 114149)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
          Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
          Pate,D. and Hood,L.
TITLE Sequencing of human chromosome 15 D15S146-D15S117 region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 114149)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
          Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
          James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
          and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2000) Multimegabase Sequencing Center, University
          of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 114149)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
          Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
          Pate,D. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-2001) Multimegabase Sequencing Center, Institute
          for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
          98105, USA
COMMENT On May 12, 2001 this sequence version replaced gi:13621222.
          ----- Genome Center
          Center: Multimegabase Sequencing Center
          Center code: UWMSC
          Web site: http://chroma.mbt.washington.edu/msg_www
          Contact: leorowen@systemsbiology.org
          ----- Summary Statistics
          Sequencing vector: pUC18; L08752
          Chemistry: Dye-terminator Big Dye; 90% of reads
          Chemistry: Dye-primer Big Dye; 10% of reads
          Assembly program: Phrap; version 0.990399

```

```

Note: Data from overlapping clones AC013355 [Drafting center:
UWMSC], AC012674 [Drafting center: BCM] and AC012378 [Drafting
center: UWMSC] were added for finishing.
FEATURES     source
             1..114149
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="15"
                /map="15q21.3"
                /clone="CTD-2137J4"
                /clone_lib="Cal Tech human BAC library D"
                /note="This clone overlaps RP11-178D12 AC013355,
                RP11-458H3 AC012674 and RP11-420M1 AC012378. Data from
                overlapping BACs were added and the consensus sequence
                determined from CTD-2137J4 to the extent possible."
             1..47896
                /note="overlap with RP11-178D12, AC013355"
             18530..18640
                misc_feature
                misc_feature

```

```

/note="sequence data generated from subcloned PCR product"
54340..54368
/note="low quality data"
90483..90630
/note="low quality data"
95860..114149
/note="overlap with RP11-458H3, AC012674"
102476..114149
/note="overlap with RP11-420M1, AC012378"

ORIGIN
Query Match          19.1%; Score 241; DB 9; Length 114149;
Best Local Similarity 92.7%; Pred. No. 3.2e-37;
Matches 253; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 398 TGATGAGATTGAAGAAGAGAGAGGAAAAAATAGAAGATATGAAGAATAATGAACGGA 457
Db |||
31582 TGTATAGATTGAAGAAGAGAGAGGAAAAAATAGAAGATATGAAGAATAATGAACGGA 31523

Qy 458 TAAAGCCACTTAAGCATTGGAGCTGGAAGAATATCAAGAATAATCAAGAAAGCTGAGGACAAA 517
Db |||
31522 TAAAGCCACTTAAGCATTGGAGCTGGAAGAATATCAAGAATAATCAAGAAAGCTGAGGACAAA 31463

Qy 518 AAAAATTTCAGAGAGAGAGAAATATTGTCAAAAAGAAAGCAAAATTAAGAGGAGAA 577
Db |||
31462 AAAAATTTCAGAGAGAGAGAAATATTGTCAAAAAGAAAGCAAAATTAAGAGGAGAA 31403

Qy 578 AAAAATAAATAATAGAGTCTTACTAGAAATTTGGCATCTAGAATCTTGTCCAAAAG 637
Db |||
31402 AAAAATAAATAATAGAGTCTTACTAGAAATTTGGCATCTAGAATCTTGTCCAAAAG 31343

Qy 638 GGAGAAATTCAGAAATATATTACTGAGAAGT 670
Db |||
31342 GTAAATGATTGAAGAAATTTCTGTAGTAGTAGT 31310

```

```

RESULT 15
AC021668/c
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-13306 map 15, LOW-PASS
SEQUENCE SAMPLING.
AC021668
VERSION AC021668.4 GI:13488020
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 43886)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Beckerly,R., Bada,F.,
Anderson,S., Baldwin,J., Barna,N., Brown,A., Burkett,G., Castle,A.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Cawley,C., Collins,S.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArrellano,K., Dewar,K., Domino,M., Doyle,M., Fenster,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lechoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testfay,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

COMMENT

```

On Mar 29, 2001 this sequence version replaced gi:11612359.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1276
Center clone name: 133_O_6
-----
* NOTE: This record contains 53 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 745: contig of 745 bp in length
* 746 845: gap of 100 bp
* 846 1576: contig of 731 bp in length
* 1577 1676: gap of 100 bp
* 1677 2410: contig of 734 bp in length
* 2411 2510: gap of 100 bp
* 2511 3232: contig of 722 bp in length
* 3233 3332: gap of 100 bp
* 3333 4060: contig of 728 bp in length
* 4061 4160: gap of 100 bp
* 4161 4891: contig of 731 bp in length
* 4892 4991: gap of 100 bp
* 4992 5708: contig of 717 bp in length
* 5709 5808: gap of 100 bp
* 5809 6537: contig of 729 bp in length
* 6538 6637: gap of 100 bp
* 6638 7344: contig of 707 bp in length
* 7345 7444: gap of 100 bp
* 7445 8167: contig of 723 bp in length
* 8168 8267: gap of 100 bp
* 8268 8995: contig of 728 bp in length
* 8996 9095: gap of 100 bp
* 9096 9839: contig of 744 bp in length
* 9840 9939: gap of 100 bp
* 9940 10653: contig of 714 bp in length
* 10654 10753: gap of 100 bp
* 10754 11497: contig of 744 bp in length
* 11498 11597: gap of 100 bp
* 11598 12315: contig of 718 bp in length
* 12316 12415: gap of 100 bp
* 12416 13136: contig of 721 bp in length
* 13137 13236: gap of 100 bp
* 13237 13967: contig of 731 bp in length
* 13968 14067: gap of 100 bp
* 14068 14802: contig of 735 bp in length
* 14803 14902: gap of 100 bp
* 14903 15642: contig of 740 bp in length
* 15643 15742: gap of 100 bp
* 15743 16477: contig of 735 bp in length
* 16478 16577: gap of 100 bp
* 16578 17301: contig of 724 bp in length
* 17302 17401: gap of 100 bp
* 17402 18129: contig of 728 bp in length
* 18130 18229: gap of 100 bp
* 18230 18958: contig of 729 bp in length
* 18959 19058: gap of 100 bp
* 19059 19787: contig of 729 bp in length
* 19788 19887: gap of 100 bp
* 19888 20627: contig of 740 bp in length

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:36:17 ; Search time 755 Seconds
(without alignments)
9902.826 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgectcttcaggtagcgca.....gaacagaaactaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	1263	9 ADB16964	ADB16964 Human DYX
2	1263	100.0	1263	9 ADB16965	ADB16965 cDNA sequ
3	1253.4	99.2	1263	9 ADB16939	ADB16939 Pygmy chi
4	1251.8	99.1	1263	9 ADB16933	ADB16933 Chimpanze
5	1248.6	98.9	1263	9 ADB16935	ADB16935 Gorilla D
6	1243.8	98.5	1263	9 ADB16937	ADB16937 Orangutan
7	1049	83.1	1641	10 ADC30210	ADC30210 Human nov
8	1047.4	82.9	1559	11 ADM01890	ADM01890 Human CDN
9	840.8	66.6	1697	9 ADB16924	ADB16924 cDNA sequ
10	524	41.5	608	10 ADC32116	ADC32116 Human nov
11	369.8	29.3	488	9 ACH35463	ACH35463 Human end
12	336	26.6	458	9 ACH23091	ACH23091 Human adu
13	245.2	19.4	1383	5 AAS70018	AAS70018 DNA encod
14	239.4	19.0	49806	9 ADB16927	ADB16927 Human DYX
15	178.8	14.2	313	3 AAC26799	AAC26799 Human sec
16	162.8	12.9	164	3 AAC30498	AAC30498 Human sec
17	156.6	12.4	49939	9 ADB16928	ADB16928 Human DYX
18	97.6	7.7	14006	6 ABL33958	ABL33958 Human imm
19	96.8	7.7	16033	6 ABL33404	ABL33404 Human imm
20	96.4	7.6	700	10 ACD92384	ACD92384 Human col

21	94.6	7.5	6644	2 AAX33181	Aax33181 Base sequ
22	94.6	7.5	7372	2 AAX33182	Aax33182 Base sequ
23	94.6	7.5	7797	2 AAX33180	Aax33180 Cowpox vi
24	94.6	7.5	7996	2 AAX33184	Aax33184 Base sequ
25	93.4	7.4	447	13 ACN54792	Acn54792 Cotton an
26	93.2	7.4	529	8 ABX52180	Abx52180 Bovine ES
27	93.2	7.4	2520	12 ADJ48201	Adj48201 Maize oil
28	92.8	7.3	627	13 ACN54555	Acn54555 Cotton oil
29	92.8	7.3	6668	6 ABL33697	Ab133697 Human imm
30	92.8	7.3	17934	6 ABL33719	Ab133719 Human imm
31	92.2	7.3	556	6 ABQ36997	Abq36997 Oligonucl
32	92.2	7.3	556	6 ABQ36996	Abq36996 Oligonucl
33	92.2	7.3	16633	6 ABN79984	Abn79984 Human che
34	91.4	7.2	1000	12 ADQ62833	Adq62833 Homopoly-
35	91.4	7.2	1000	12 ADQ62832	Adq62832 Homopoly-
36	91.4	7.2	1300	12 ADP85917	Adp85917 Synthetic
37	91.2	7.2	3505	12 ADJ48200	Adj48200 Maize oil
38	89.8	7.1	2226	12 ADI43251	Adi43251 Plant tra
39	89.8	7.1	2226	12 ADO02941	Ado02941 Soybean o
40	89.6	7.1	34769	4 AAS46774	Aas46774 Tumour su
41	89.4	7.1	113515	6 ABL34174	Ab134174 Human imm
42	88.4	7.0	537	13 ACN53339	Acn53339 Cotton an
43	88.2	7.0	1179	6 ABQ49604	Abq49604 Oligonucl
44	88.2	7.0	1179	6 ABQ49605	Abq49605 Oligonucl
45	87.6	6.9	529	13 ACN56029	Acn56029 Cotton an

ALIGNMENTS

RESULT 1
ADB16964
ID ADB16964 standard; cDNA; 1263 BP.
XX
AC ADB16964;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human DYXC1 cDNA with single nucleotide polymorphisms.
XX
KW Gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW Chromosome 15q21; reading disability; phonological processing;
KW rapid naming; verbal short-term memory; single nucleotide polymorphism;
KW SNP.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 1..1283
FT /product= "DYXC1 protein"
FT /tag= a
FT /tag= b
FT /tag= c
FT /tag= d
FT /tag= e
FT /tag= f
FT /tag= g
FT /tag= h
FT /tag= i
FT /tag= j
FT /tag= k
FT /tag= l
FT /tag= m
FT /tag= n
FT /tag= o
FT /tag= p
FT /tag= q
FT /tag= r
FT /tag= s
FT /tag= t
FT /tag= u
FT /tag= v
FT /tag= w
FT /tag= x
FT /tag= y
FT /tag= z
FT /tag= AA
FT /tag= AB
FT /tag= AC
FT /tag= AD
FT /tag= AE
FT /tag= AF
FT /tag= AG
FT /tag= AH
FT /tag= AI
FT /tag= AJ
FT /tag= AK
FT /tag= AL
FT /tag= AM
FT /tag= AN
FT /tag= AO
FT /tag= AP
FT /tag= AQ
FT /tag= AR
FT /tag= AS
FT /tag= AT
FT /tag= AU
FT /tag= AV
FT /tag= AW
FT /tag= AX
FT /tag= AY
FT /tag= AZ
FT /tag= BA
FT /tag= BB
FT /tag= BC
FT /tag= BD
FT /tag= BE
FT /tag= BF
FT /tag= BG
FT /tag= BH
FT /tag= BI
FT /tag= BJ
FT /tag= BK
FT /tag= BL
FT /tag= BM
FT /tag= BN
FT /tag= BO
FT /tag= BP
FT /tag= BQ
FT /tag= BR
FT /tag= BS
FT /tag= BT
FT /tag= BU
FT /tag= BV
FT /tag= BW
FT /tag= BX
FT /tag= BY
FT /tag= BZ
FT /tag= CA
FT /tag= CB
FT /tag= CC
FT /tag= CD
FT /tag= CE
FT /tag= CF
FT /tag= CG
FT /tag= CH
FT /tag= CI
FT /tag= CJ
FT /tag= CK
FT /tag= CL
FT /tag= CM
FT /tag= CN
FT /tag= CO
FT /tag= CP
FT /tag= CQ
FT /tag= CR
FT /tag= CS
FT /tag= CT
FT /tag= CU
FT /tag= CV
FT /tag= CW
FT /tag= CX
FT /tag= CY
FT /tag= CZ
FT /tag= DA
FT /tag= DB
FT /tag= DC
FT /tag= DD
FT /tag= DE
FT /tag= DF
FT /tag= DG
FT /tag= DH
FT /tag= DI
FT /tag= DJ
FT /tag= DK
FT /tag= DL
FT /tag= DM
FT /tag= DN
FT /tag= DO
FT /tag= DP
FT /tag= DQ
FT /tag= DR
FT /tag= DS
FT /tag= DT
FT /tag= DU
FT /tag= DV
FT /tag= DW
FT /tag= DX
FT /tag= DY
FT /tag= DZ
FT /tag= EA
FT /tag= EB
FT /tag= EC
FT /tag= ED
FT /tag= EE
FT /tag= EF
FT /tag= EG
FT /tag= EH
FT /tag= EI
FT /tag= EJ
FT /tag= EK
FT /tag= EL
FT /tag= EM
FT /tag= EN
FT /tag= EO
FT /tag= EP
FT /tag= EQ
FT /tag= ER
FT /tag= ES
FT /tag= ET
FT /tag= EU
FT /tag= EV
FT /tag= EW
FT /tag= EX
FT /tag= EY
FT /tag= EZ
FT /tag= FA
FT /tag= FB
FT /tag= FC
FT /tag= FD
FT /tag= FE
FT /tag= FF
FT /tag= FG
FT /tag= FH
FT /tag= FI
FT /tag= FJ
FT /tag= FK
FT /tag= FL
FT /tag= FM
FT /tag= FN
FT /tag= FO
FT /tag= FP
FT /tag= FQ
FT /tag= FR
FT /tag= FS
FT /tag= FT
FT /tag= FU
FT /tag= FV
FT /tag= FW
FT /tag= FX
FT /tag= FY
FT /tag= FZ
FT /tag= GA
FT /tag= GB
FT /tag= GC
FT /tag= GD
FT /tag= GE
FT /tag= GF
FT /tag= GG
FT /tag= GH
FT /tag= GI
FT /tag= GJ
FT /tag= GK
FT /tag= GL
FT /tag= GM
FT /tag= GN
FT /tag= GO
FT /tag= GP
FT /tag= GQ
FT /tag= GR
FT /tag= GS
FT /tag= GT
FT /tag= GU
FT /tag= GV
FT /tag= GW
FT /tag= GX
FT /tag= GY
FT /tag= GZ
FT /tag= HA
FT /tag= HB
FT /tag= HC
FT /tag= HD
FT /tag= HE
FT /tag= HF
FT /tag= HG
FT /tag= HH
FT /tag= HI
FT /tag= HJ
FT /tag= HK
FT /tag= HL
FT /tag= HM
FT /tag= HN
FT /tag= HO
FT /tag= HP
FT /tag= HQ
FT /tag= HR
FT /tag= HS
FT /tag= HT
FT /tag= HU
FT /tag= HV
FT /tag= HW
FT /tag= HX
FT /tag= HY
FT /tag= HZ
FT /tag= IA
FT /tag= IB
FT /tag= IC
FT /tag= ID
FT /tag= IE
FT /tag= IF
FT /tag= IG
FT /tag= IH
FT /tag= II
FT /tag= IJ
FT /tag= IK
FT /tag= IL
FT /tag= IM
FT /tag= IN
FT /tag= IO
FT /tag= IP
FT /tag= IQ
FT /tag= IR
FT /tag= IS
FT /tag= IT
FT /tag= IU
FT /tag= IV
FT /tag= IW
FT /tag= IX
FT /tag= IY
FT /tag= IZ
FT /tag= JA
FT /tag= JB
FT /tag= JC
FT /tag= JD
FT /tag= JE
FT /tag= JF
FT /tag= JG
FT /tag= JH
FT /tag= JI
FT /tag= JJ
FT /tag= JK
FT /tag= JL
FT /tag= JM
FT /tag= JN
FT /tag= JO
FT /tag= JP
FT /tag= JQ
FT /tag= JR
FT /tag= JS
FT /tag= JT
FT /tag= JU
FT /tag= JV
FT /tag= JW
FT /tag= JX
FT /tag= JY
FT /tag= JZ
FT /tag= KA
FT /tag= KB
FT /tag= KC
FT /tag= KD
FT /tag= KE
FT /tag= KF
FT /tag= KG
FT /tag= KH
FT /tag= KI
FT /tag= KJ
FT /tag= KK
FT /tag= KL
FT /tag= KM
FT /tag= KN
FT /tag= KO
FT /tag= KP
FT /tag= KQ
FT /tag= KR
FT /tag= KS
FT /tag= KT
FT /tag= KU
FT /tag= KV
FT /tag= KW
FT /tag= KX
FT /tag= KY
FT /tag= KZ
FT /tag= LA
FT /tag= LB
FT /tag= LC
FT /tag= LD
FT /tag= LE
FT /tag= LF
FT /tag= LG
FT /tag= LH
FT /tag= LI
FT /tag= LJ
FT /tag= LK
FT /tag= LL
FT /tag= LM
FT /tag= LN
FT /tag= LO
FT /tag= LP
FT /tag= LQ
FT /tag= LR
FT /tag= LS
FT /tag= LT
FT /tag= LU
FT /tag= LV
FT /tag= LW
FT /tag= LX
FT /tag= LY
FT /tag= LZ
FT /tag= MA
FT /tag= MB
FT /tag= MC
FT /tag= MD
FT /tag= ME
FT /tag= MF
FT /tag= MG
FT /tag= MH
FT /tag= MI
FT /tag= MJ
FT /tag= MK
FT /tag= ML
FT /tag= MN
FT /tag= MO
FT /tag= MP
FT /tag= MQ
FT /tag= MR
FT /tag= MS
FT /tag= MT
FT /tag= MU
FT /tag= MV
FT /tag= MW
FT /tag= MX
FT /tag= MY
FT /tag= MZ
FT /tag= NA
FT /tag= NB
FT /tag= NC
FT /tag= ND
FT /tag= NE
FT /tag= NF
FT /tag= NG
FT /tag= NH
FT /tag= NI
FT /tag= NJ
FT /tag= NK
FT /tag= NL
FT /tag= NM
FT /tag= NO
FT /tag= NP
FT /tag= NQ
FT /tag= NR
FT /tag= NS
FT /tag= NT
FT /tag= NU
FT /tag= NV
FT /tag= NW
FT /tag= NX
FT /tag= NY
FT /tag= NZ
FT /tag= OA
FT /tag= OB
FT /tag= OC
FT /tag= OD
FT /tag= OE
FT /tag= OF
FT /tag= OG
FT /tag= OH
FT /tag= OI
FT /tag= OJ
FT /tag= OK
FT /tag= OL
FT /tag= OM
FT /tag= ON
FT /tag= OO
FT /tag= OP
FT /tag= OQ
FT /tag= OR
FT /tag= OS
FT /tag= OT
FT /tag= OU
FT /tag= OV
FT /tag= OW
FT /tag= OX
FT /tag= OY
FT /tag= OZ
FT /tag= PA
FT /tag= PB
FT /tag= PC
FT /tag= PD
FT /tag= PE
FT /tag= PF
FT /tag= PG
FT /tag= PH
FT /tag= PI
FT /tag= PJ
FT /tag= PK
FT /tag= PL
FT /tag= PM
FT /tag= PN
FT /tag= PO
FT /tag= PP
FT /tag= PQ
FT /tag= PR
FT /tag= PS
FT /tag= PT
FT /tag= PU
FT /tag= PV
FT /tag= PW
FT /tag= PX
FT /tag= PY
FT /tag= PZ
FT /tag= QA
FT /tag= QB
FT /tag= QC
FT /tag= QD
FT /tag= QE
FT /tag= QF
FT /tag= QG
FT /tag= QH
FT /tag= QI
FT /tag= QJ
FT /tag= QK
FT /tag= QL
FT /tag= QM
FT /tag= QN
FT /tag= QO
FT /tag= QP
FT /tag= QQ
FT /tag= QR
FT /tag= QS
FT /tag= QT
FT /tag= QU
FT /tag= QV
FT /tag= QW
FT /tag= QX
FT /tag= QY
FT /tag= QZ
FT /tag= RA
FT /tag= RB
FT /tag= RC
FT /tag= RD
FT /tag= RE
FT /tag= RF
FT /tag= RG
FT /tag= RH
FT /tag= RI
FT /tag= RJ
FT /tag= RK
FT /tag= RL
FT /tag= RM
FT /tag= RN
FT /tag= RO
FT /tag= RP
FT /tag= RQ
FT /tag= RR
FT /tag= RS
FT /tag= RT
FT /tag= RU
FT /tag= RV
FT /tag= RW
FT /tag= RX
FT /tag= RY
FT /tag= RZ
FT /tag= SA
FT /tag= SB
FT /tag= SC
FT /tag= SD
FT /tag= SE
FT /tag= SF
FT /tag= SG
FT /tag= SH
FT /tag= SI
FT /tag= SJ
FT /tag= SK
FT /tag= SL
FT /tag= SM
FT /tag= SN
FT /tag= SO
FT /tag= SP
FT /tag= SQ
FT /tag= SR
FT /tag= SS
FT /tag= ST
FT /tag= SU
FT /tag= SV
FT /tag= SW
FT /tag= SX
FT /tag= SY
FT /tag= SZ
FT /tag= TA
FT /tag= TB
FT /tag= TC
FT /tag= TD
FT /tag= TE
FT /tag= TF
FT /tag= TG
FT /tag= TH
FT /tag= TI
FT /tag= TJ
FT /tag= TK
FT /tag= TL
FT /tag= TM
FT /tag= TN
FT /tag= TO
FT /tag= TP
FT /tag= TQ
FT /tag= TR
FT /tag= TS
FT /tag= TT
FT /tag= TU
FT /tag= TV
FT /tag= TW
FT /tag= TX
FT /tag= TY
FT /tag= TZ
FT /tag= UA
FT /tag= UB
FT /tag= UC
FT /tag= UD
FT /tag= UE
FT /tag= UF
FT /tag= UG
FT /tag= UH
FT /tag= UI
FT /tag= UJ
FT /tag= UK
FT /tag= UL
FT /tag= UM
FT /tag= UN
FT /tag= UO
FT /tag= UP
FT /tag= UQ
FT /tag= UR
FT /tag= US
FT /tag= UT
FT /tag= UU
FT /tag= UV
FT /tag= UW
FT /tag= UX
FT /tag= UY
FT /tag= UZ
FT /tag= VA
FT /tag= VB
FT /tag= VC
FT /tag= VD
FT /tag= VE
FT /tag= VF
FT /tag= VG
FT /tag= VH
FT /tag= VI
FT /tag= VJ
FT /tag= VK
FT /tag= VL
FT /tag= VM
FT /tag= VN
FT /tag= VO
FT /tag= VP
FT /tag= VQ
FT /tag= VR
FT /tag= VS
FT /tag= VT
FT /tag= VU
FT /tag= VV
FT /tag= VW
FT /tag= VX
FT /tag= VY
FT /tag= VZ
FT /tag= WA
FT /tag= WB
FT /tag= WC
FT /tag= WD
FT /tag= WE
FT /tag= WF
FT /tag= WG
FT /tag= WH
FT /tag= WI
FT /tag= WJ
FT /tag= WK
FT /tag= WL
FT /tag= WM
FT /tag= WN
FT /tag= WO
FT /tag= WP
FT /tag= WQ
FT /tag= WR
FT /tag= WS
FT /tag= WT
FT /tag= WU
FT /tag= WV
FT /tag= WW
FT /tag= WX
FT /tag= WY
FT /tag= WZ
FT /tag= XA
FT /tag= XB
FT /tag= XC
FT /tag= XD
FT /tag= XE
FT /tag= XF
FT /tag= XG
FT /tag= XH
FT /tag= XI
FT /tag= XJ
FT /tag= XK
FT /tag= XL
FT /tag= XM
FT /tag= XN
FT /tag= XO
FT /tag= XP
FT /tag= XQ
FT /tag= XR
FT /tag= XS
FT /tag= XT
FT /tag= XU
FT /tag= XV
FT /tag= XW
FT /tag= XX
FT /tag= XY
FT /tag= XZ
FT /tag= YA
FT /tag= YB
FT /tag= YC
FT /tag= YD
FT /tag= YE
FT /tag= YF
FT /tag= YG
FT /tag= YH
FT /tag= YI
FT /tag= YJ
FT /tag= YK
FT /tag= YL
FT /tag= YM
FT /tag= YN
FT /tag= YO
FT /tag= YP
FT /tag= YQ
FT /tag= YR
FT /tag= YS
FT /tag= YT
FT /tag= YU
FT /tag= YV
FT /tag= YW
FT /tag= YX
FT /tag= YY
FT /tag= YZ
FT /tag= ZA
FT /tag= ZB
FT /tag= ZC
FT /tag= ZD
FT /tag= ZE
FT /tag= ZF
FT /tag= ZG
FT /tag= ZH
FT /tag= ZI
FT /tag= ZJ
FT /tag= ZK
FT /tag= ZL
FT /tag= ZM
FT /tag= ZN
FT /tag= ZO
FT /tag= ZP
FT /tag= ZQ
FT /tag= ZR
FT /tag= ZS
FT /tag= ZT
FT /tag= ZU
FT /tag= ZV
FT /tag= ZW
FT /tag= ZX
FT /tag= ZY
FT /tag= ZZ
FT /tag= AA
FT /tag= AB
FT /tag= AC
FT /tag= AD
FT /tag= AE
FT /tag= AF
FT /tag= AG
FT /tag= AH
FT /tag= AI
FT /tag= AJ
FT /tag= AK
FT /tag= AL
FT /tag= AM
FT /tag= AN
FT /tag= AO
FT /tag= AP
FT /tag= AQ
FT /tag= AR
FT /tag= AS
FT /tag= AT
FT /tag= AU
FT /tag= AV
FT /tag= AW
FT /tag= AX
FT /tag= AY
FT /tag= AZ
FT /tag= BA
FT /tag= BB
FT /tag= BC
FT /tag= BD
FT /tag= BE
FT /tag= BF
FT /tag= BG
FT /tag= BH
FT /tag= BI
FT /tag= BJ
FT /tag= BK
FT /tag= BL
FT /tag= BM
FT /tag= BN
FT /tag= BO
FT /tag= BP
FT /tag= BQ
FT /tag= BR
FT /tag= BS
FT /tag= BT
FT /tag= BU
FT /tag= BV
FT /tag= BW
FT /tag= BX
FT /tag= BY
FT /tag= BZ
FT /tag= CA
FT /tag= CB
FT /tag= CC
FT /tag= CD
FT /tag= CE
FT /tag= CF
FT /tag= CG
FT /tag= CH
FT /tag= CI
FT /tag= CJ
FT /tag= CK
FT /tag= CL
FT /tag= CM
FT /tag= CN
FT /tag= CO
FT /tag= CP
FT /tag= CQ
FT /tag= CR
FT /tag= CS
FT /tag= CT
FT /tag= CU
FT /tag= CV
FT /tag= CW
FT /tag= CX
FT /tag= CY
FT /tag= CZ
FT /tag= DA
FT /tag= DB
FT /tag= DC
FT /tag= DD
FT /tag= DE
FT /tag= DF
FT /tag= DG
FT /tag= DH
FT /tag= DI
FT /tag= DJ
FT /tag= DK
FT /tag= DL
FT /tag= DM
FT /tag= DN
FT /tag= DO
FT /tag= DP
FT /tag= DQ
FT /tag= DR
FT /tag= DS
FT /tag= DT
FT /tag= DU
FT /tag= DV
FT /tag= DW
FT /tag= DX
FT /tag= DY
FT /tag= DZ
FT /tag= EA
FT /tag= EB
FT /tag= EC
FT /tag= ED
FT /tag= EE
FT /tag= EF
FT /tag= EG
FT /tag= EH
FT /tag= EI
FT /tag= EJ
FT /tag= EK
FT /tag= EL
FT /tag= EM
FT /tag= EN
FT /tag= EO
FT /tag= EP
FT /tag= EQ
FT /tag= ER
FT /tag= ES
FT /tag= ET
FT /tag= EU
FT /tag= EV
FT /tag= EW
FT /tag= EX
FT /tag= EY
FT /tag= EZ
FT /tag= FA
FT /tag= FB
FT /tag= FC
FT /tag= FD
FT /tag= FE
FT /tag= FG
FT /tag= FH
FT /tag= FI
FT /tag= FJ
FT /tag= FK
FT /tag= FL
FT /tag= FM
FT /tag= FN
FT /tag= FO
FT /tag= FP
FT /tag= FQ
FT /tag= FR
FT /tag= FS
FT /tag= FT
FT /tag= FU
FT /tag= FV
FT /tag= FW
FT /tag= FX
FT /tag= FY
FT /tag= FZ
FT /tag= GA
FT /tag= GB
FT /tag= GC
FT /tag= GD
FT /tag= GE
FT /tag= GF
FT /tag= GG
FT /tag= GH
FT /tag= GI
FT /tag= GJ
FT /tag= GK
FT /tag= GL
FT /tag= GM
FT /tag= GN
FT /tag= GO
FT /tag= GP
FT /tag= GQ
FT /tag= GR
FT /tag= GS
FT /tag= GT
FT /tag= GU
FT /tag= GV
FT /tag= GW
FT /tag= GX
FT /tag= GY
FT /tag= GZ
FT /tag= HA
FT /tag= HB
FT /tag= HC
FT /tag= HD
FT /tag= HE
FT /tag= HF
FT /tag= HG
FT /tag= HH
FT /tag= HI
FT /tag= HJ
FT /tag= HK
FT /tag= HL
FT /tag= HM
FT /tag= HN
FT /tag= HO
FT /tag= HP
FT /tag= HQ
FT /tag= HR
FT /tag= HS
FT /tag= HT
FT /tag= HU
FT /tag= HV
FT /tag= HW
FT /tag= HX
FT /tag= HY
FT /tag= HZ
FT /tag= IA
FT /tag= IB
FT /tag= IC
FT /tag= ID
FT /tag= IE
FT /tag= IF
FT /tag= IG
FT /tag= IH
FT /tag= II
FT /tag= IJ
FT /tag= IK
FT /tag= IL
FT /tag= IM
FT /tag= IN
FT /tag= IO
FT /tag= IP
FT /tag= IQ
FT /tag= IR
FT /tag= IS
FT /tag= IT
FT /tag= IU
FT /tag= IV
FT /tag= IW
FT /tag= IX
FT /tag= IY
FT /tag= IZ
FT /tag= JA
FT /tag= JB
FT /tag= JC
FT /tag= JD
FT /tag= JE
FT /tag= JF
FT /tag= JG
FT /tag= JH
FT /tag= JI
FT /tag= JJ
FT /tag= JK
FT /tag= JL
FT /tag= JM
FT /tag= JN
FT /tag= JO
FT /tag= JP
FT /tag= JQ
FT /tag= JR
FT /tag= JS
FT /tag= JT
FT /tag= JU
FT /tag= JV
FT /tag= JW
FT /tag= JX
FT /tag= JY
FT /tag= JZ
FT /tag= KA
FT /tag= KB
FT /tag= KC
FT /tag= KD
FT /tag= KE
FT /tag= KF
FT /tag= KG
FT /tag= KH
FT /tag= KI
FT /tag= KJ
FT /tag= KK
FT /tag= KL
FT /tag= KM
FT /tag= KN
FT /tag= KO
FT /tag= KP
FT /tag= KQ
FT /tag= KR
FT /tag= KS
FT /tag= KT
FT /tag= KU
FT /tag= KV
FT /tag= KW
FT /tag= KX
FT /tag= KY
FT /tag= KZ
FT /tag= LA
FT /tag= LB
FT /tag= LC
FT /tag= LD
FT /tag= LE
FT /tag= LF
FT /tag= LG
FT /tag= LH
FT /tag= LI
FT /tag= LJ
FT /tag= LK
FT /tag= LL
FT /tag= LM
FT /tag= LN
FT /tag= LO
FT /tag= LP
FT /tag= LQ
FT /tag= LR
FT /tag= LS
FT /tag= LT
FT /tag= LU
FT /tag= LV
FT /tag= LW
FT /tag= LX
FT /tag= LY
FT /tag= LZ
FT /tag= MA
FT /tag= MB
FT /tag= MC
FT /tag= MD
FT /tag= ME
FT /tag= MF
FT /tag= MG
FT /tag= MH
FT /tag= MI
FT /tag= MJ
FT /tag= MK
FT /tag= ML
FT /tag= MN
FT /tag= MO
FT /tag= MP
FT /tag= MQ
FT /tag= MR
FT /tag= MS
FT /tag= MT
FT /tag= MU
FT /tag= MV
FT /tag= MW
FT /tag= MX
FT /tag= MY
FT /tag= MZ
FT /tag= NA
FT /tag= NB
FT /tag= NC
FT /tag= ND
FT /tag= NE
FT /tag= NF
FT /tag= NG
FT /tag= NH
FT /tag= NI
FT /tag= NJ
FT /tag= NK
FT /tag= NL
FT /tag= NM
FT /tag= NO
FT /tag= NP
FT /tag= NQ
FT /tag= NR
FT /tag= NS
FT /tag= NT
FT /tag= NU
FT /tag= NV
FT /tag= NW
FT /tag= NX
FT /tag= NY
FT /tag= NZ
FT /tag= OA
FT /tag= OB
FT /tag= OC
FT /tag= OD
FT /tag= OE
FT /tag= OF
FT /tag= OG
FT /tag= OH
FT /tag= OI
FT /tag= OJ
FT /tag= OK
FT /tag= OL
FT /tag= OM
FT /tag= ON
FT /tag= OO
FT /tag= OP
FT /tag= OQ
FT /tag= OR
FT /tag= OS
FT /tag= OT
FT /tag= OU
FT /tag= OV
FT /tag= OW
FT /tag= OX
FT /tag= OY
FT /tag= OZ
FT /tag= PA
FT /tag= PB
FT /tag= PC
FT /tag= PD
FT /tag= PE
FT /tag= PF
FT /tag= PG
FT /tag= PH
FT /tag= PI
FT /tag= PJ
FT /tag= PK
FT /tag= PL
FT /tag= PM
FT /tag= PN
FT /tag= PO
FT /tag= PP
FT /tag= PQ
FT /tag= PR
FT /tag= PS
FT /tag= PT
FT /tag= PU
FT /tag= PV
FT

XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX WPI; 2003-646482/61.
DR P-PSDB; ADB16923.
XX
XX New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
XX Claim 1; Page 47; 135pp; English.
XX
XX This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide
CC represents the coding sequence of human DYXC1 cDNA of the invention.
XX
XX Sequence 1263 BP; 491 A; 214 C; 275 G; 283 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1263; DB 9; Length 1263;
Best Local Similarity 100.0%; Pred. No. 1.3e-263;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCTCTTCAGGTACGATTAACAGCTGGCAGCAGACGAGACTGCGTCTTCTGTCT 60
Db 1 ATGCCTCTTCAGGTACGATTAACAGCTGGCAGCAGACGAGACTGCGTCTTCTGTCT 60
Qy 61 CTGCCCCCAAGCGGTGCGTCAGAGACGCGCTTCTGACGGAACATCTCTG 120
Db 61 CTGCCCCCAAGCGGTGCGTCAGAGACGCGCTTCTGACGGAACATCTCTG 120
Qy 121 AAGGTCAACTTCTCCATTTTATTTAGGCGATTTCTTATGTCCTCCATAGCGATGAG 180
Db 121 AAGGTCAACTTCTCCATTTTATTTAGGCGATTTCTTATGTCCTCCATAGCGATGAG 180
Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCCTTGTATATAAAGAGCG 240
Db 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCCTTGTATATAAAGAGCG 240
Qy 241 GCCATGTGGGAGACCCCTTCTGTGACCGGTGTCACAAAGAGATGTCGAAAGATTAGA 300
Db 241 GCCATGTGGGAGACCCCTTCTGTGACCGGTGTCACAAAGAGATGTCGAAAGATTAGA 300
Qy 301 GAAAAATCTATTTTACAGCACAAGAGAGACCAAGCAAGCTACAGAAAGCTGCA 360
Db 301 GAAAAATCTATTTTACAGCACAAGAGAGACCAAGCAAGCTACAGAAAGCTGCA 360
Qy 361 GCAAGCGGGAAGATCAAAAATACGCACTAAGTGTGATGATGAAGATTGAAGAGAGAG 420
Db 361 GCAAGCGGGAAGATCAAAAATACGCACTAAGTGTGATGATGAAGATTGAAGAGAGAG 420
Qy 421 AGGAAAAATAGAGATGAAAGAAATGAACCGGATTAAGCGCTTAAAGCATTGGA 480
Db 421 AGGAAAAATAGAGATGAAAGAAATGAACCGGATTAAGCGCTTAAAGCATTGGA 480
Qy 481 GCCTGGAAGAAATCAAGAAAAAGCTGAGGACAAAAAATTCAGAGAGAGAGAAA 540
Db 481 GCCTGGAAGAAATCAAGAAAAAGCTGAGGACAAAAAATTCAGAGAGAGAGAAA 540
Qy 541 TTATGTCAAAAGAAAAAGCAATTAAGAGAGAGAAAAAATAAATAAAGAGTCTT 600
Db 541 TTATGTCAAAAGAAAAAGCAATTAAGAGAGAGAAAAAATAAATAAAGAGTCTT 600
Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGAGAAATTCAGAAAAATATTT 660

Db 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGAGAAATTCAGAAAAATATATTT 660
Qy 661 ACTGAGAGAGTTAAAGGAGACAGTATTCCTGCTCTCTGCTCTGTTGGCAGTATTAATTC 720
Db 661 ACTGAGAGAGTTAAAGGAGACAGTATTCCTGCTCTCTGCTCTGTTGGCAGTATTAATTC 720
Qy 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTAATCAACAAGTAGCAGAGAGAG 780
Db 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTAATCAACAAGTAGCAGAGAGAG 780
Qy 781 GAGTGGCTACACAAAAGCTGAGGACCAAGAGCAATGAATGATCTGACATAGCTGAATTT 840
Db 781 GAGTGGCTACACAAAAGCTGAGGACCAAGAGCAATGAATGATCTGACATAGCTGAATTT 840
Qy 841 TGCATTTAAAGAGAGAGAAAAAGACCCAGATGTTGAAGATTAAGAGAGAGAGAGAGAG 900
Db 841 TGCATTTAAAGAGAGAGAAAAAGACCCAGATGTTGAAGATTAAGAGAGAGAGAGAGAG 900
Qy 901 TTTGCAACGGAACCTATTTGGCAGCTATCAATGCATATTAATTTAGCCATAAGACTAAAT 960
Db 901 TTTGCAACGGAACCTATTTGGCAGCTATCAATGCATATTAATTTAGCCATAAGACTAAAT 960
Qy 961 AATAAGATGCACTATTTGTAATTTGAACCGGCTGCTTGCACCTAAAAAATCTTA 1020
Db 961 AATAAGATGCACTATTTGTAATTTGAACCGGCTGCTTGCACCTAAAAAATCTTA 1020
Qy 1021 CACAAGCTATTGAAGATTTCTTAAGGCACTGGAATTTATGATGCCACCTGTTTACAGAC 1080
Db 1021 CACAAGCTATTGAAGATTTCTTAAGGCACTGGAATTTATGATGCCACCTGTTTACAGAC 1080
Qy 1081 AATGCTAATCAAGATTAAGGACACATGTACGACGTCGGAACAGCATTTCTGCAACTAGAA 1140
Db 1081 AATGCTAATCAAGATTAAGGACACATGTACGACGTCGGAACAGCATTTCTGCAACTAGAA 1140
Qy 1141 TTGATGTAGAGCGCTTACAGGATTTATGAACGCGCACTTAAGATTTGATCCATCCACAAA 1200
Db 1141 TTGATGTAGAGCGCTTACAGGATTTATGAACGCGCACTTAAGATTTGATCCATCCACAAA 1200
Qy 1201 ATTGTACAAATGATGCTGAGAGATTCGGAATGTAATTAAGAGAGAGAGAGAGAGAG 1260
Db 1201 ATTGTACAAATGATGCTGAGAGATTCGGAATGTAATTAAGAGAGAGAGAGAGAGAG 1260
Qy 1261 TAA 1263
Db 1261 TAA 1263
RESULT 2
ADBI6965
ID ADBI6965 standard; cDNA; 1993 BP.
XX
AC ADBI6965;
XX AC
XX DT 20-NOV-2003 (first entry)
XX DE cDNA sequence of the human DYXC1 mRNA.
XX KW gene; ss; human; DYXC1; dyslexia; neurological disorder;
XX KW chromosome 15q21; reading disability; phonological processing;
XX KW rapid naming; verbal short-term memory; SNP;
XX KW single nucleotide polymorphism.
XX OS Homo sapiens.
XX XX
XX Key Location/Qualifiers
XX variation replace (205,t)
XX /*tag= a
XX /standard_name= "Single nucleotide polymorphism"
XX CDS 369..1631
XX /*tag= b
XX /product= "DYXC1 protein"
XX XX

gene; ss: chimpanzee; DYXC1; dyslexia; neurological disorder; reading disability; phonological processing; rapid naming; verbal short-term memory.

Pan troglodytes.

Key	Location/Qualifiers
CDS	1. .1263

CDS
1. .1263

```
/*tag= a
/product= "DYXC1 protein"
```

```

/product= "DYXC1 protein"

```

WO2003068814-A1.

21-AUG-2003.

12-FEB-2003: 2003W0-ET000110

13-PPB-3003, 2002118-03557928

12 MAY 1964

Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;

WPI: 2003-646482/61.

P-PSDB: ADB16934.

New isolated, purified DVXCl nucleic acid for studying brain processes, e.g. reading, phonological processing, rapid naming or verbal short-term memory, or for diagnosing dyslexia or assessing the predisposition to dyslexia.

Claim 29: Page 116-118: 135pp: English.

This invention relates to a novel isolated human gene *DYX1C1* that is functionally related to dyslexia, more particularly it describes single nucleotide polymorphisms thought to predispose an individual in to developing dyslexia. This is a neurological disorder with a genetic basis (*DYX1C1* has been isolated to chromosome 15q21), which manifests itself as a specific reading disability. Specifically, *DYX1C1* is can be useful in a study of brain processes such as reading, phonological processing, rapid naming and verbal short-term memory. Accordingly, the present invention describes methods and materials for analysing allelic variations in the *DYX1C1* gene, and also provides *DYX1C1* as an antigen for the production of antibodies used in the diagnosis of dyslexia. This polynucleotide sequence is the chimpanzee *DYX1C1* cDNA homologous to the human *DYX1C1* gene of the invention.

Sequence 1263 BP: 494 A: 213 C: 272 G: 284 T: 0 U: 0 Other:

any Match	99 1%:	Score 1251	8:	DB 9:	Length 1263:
-----------	--------	------------	----	-------	--------------

Only Match
but Local Similarity
99.4%; Pred No 3 4e-261;
59.1%; SCORE 1251.8; DB 5;

Local Similarity 99.4%; PRed: NO: 3.4E-201;
 Mismatches 0: Mismatches 7: Indels 0: Gaps 0:

1 ATGCGTCTTCAGGTTAGCGGATTACAGCTGGCAGCAGACGAGACTGGCGGTCTTTCTGTCT 60

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

;

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

RESULT 5

ADB16935

ADB16935
ID ADB16935 standard: cDNA: 1263 BP.

XXI

301	GAATAATCTATTTTACAGACACAGAGAGAGCAAAAGAGAGCTACAGAGC	360	AAAGCTGCA
301	GAATAATCTATTTTACAGACACAGAGAGAGCAAAAGAGAGCTACAGAGC	360	AAAGCTGCA
361	GCAAAGCGGGAAGATCAAAATAACGACCTAAGCTGTCATGATGAGATTGAAGAAGAAGAG	420	
361	GCAAAGCGGGAAGATCAAAATAATGCACTAAGCTCATGATGAGATTGAAGAAGAAGAG	420	
421	AGGAAAAAATAGAAGATATGAAGAAAAATGAACGGATATAAAGCCACTAAAGCAATTGGAA	480	
421	AGGAAAAAATAGAAGATATGAAGAAAAATGAACGGATATAAAGCCACTAAAGCAATTGGAA	480	
481	GCCTGGAAAGAATATCAAGAAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA	540	
481	GCCTGGAAAGAATATCAAGAAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA	540	
541	TTATGTCAAAAAGAAAAGCAATTAAGAAAGAGAAAAAAAATAAAATATAAGAGTCTT	600	
541	TTATGTCAAAAAGAAAAGCAATTAAGAAAGAGAAAAAAAATAAAATATAAGAGTCTT	600	
601	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAAAATTCAGAAAAATATATTT	660	
601	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAAAATTCAGAAAAATATATTT	660	
661	ACTGAGAAAGTTAAAGGAGACAGTATCTGCTCCCTGCTGTTGGCAGTATTAATAATC	720	
661	ACTGAGAAAGTTAAAGGAGACAGTATCTGCTCCCTGCTGTTGGCAGTATTAATAATC	720	
721	AACCTTTACCCCTCGAGTATTCCTCCACAGCTCTTCGTGAAATCACAAGTAGCAGAGAGGAG	780	
721	AACCTTTACCCCTCGAGTATTCCTCCACAGCTCTTCGTGAAATCACAAGTAGCAGAGAGGAG	780	
781	GAGTGGCTACACAAAACAGCTGAGGACGAAAGAGCAATGAATATCTGACATAGCTGAACCTT	840	
781	GAGTGGCTACACAAAACAGCTGAGGACGAAAGAGCAATGAATATCTGACATAGCTGAACCTT	840	
841	TGCGATTTAAAGAGAGAAAGAACCCAGAAATGGTTGAAGGATTAAGGAGAAACAAATTTG	900	
841	TGCGATTTAAAGAGAGAAAGAACCCAGAAATGGTTGAAGGATTAAGGAGAAACAAATTTG	900	
901	TTTGCAACGAAAACCTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT	960	
901	TTTGCAACGAAAACCTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT	960	
961	AATAAGATGCCACTATTGTATTGTAACCGGGCTGCTGCCACCTAAAAAATAAAAACTTA	1020	
961	AATAAGATGCCACTATTGTATTGTAACCGGGCTGCTGCCACCTAAAAAATAAAAACTTA	1020	
1021	CACAAGCTATTGAAGATCTTTCTAAGGCACTGGAAATTTATGATGCCACCTGTTACAGAC	1080	
1021	CACAAGCTATTGAAGATCTTTCTAAGGCACTGGAAATTTATGATGCCACCTGTTACAGAC	1080	
1081	AATGCTTAATGCAAGATGAAGGACACATGTACGACGTGGAAACAGCAATCTGTCAACTAGAA	1140	
1081	AATGCTTAATGCAAGATGAAGGACACATGTACGACGTGGAAACAGCAATCTGTCAACTAGAA	1140	
1141	TTGTATCTAGAGGCTTACAGGATTAAGCGGCACCTTAAGATTGATCCATCCAAACAA	1200	
1141	TTGTATCTAGAGGCTTACAGGATTAAGCGGCACCTTAAGATTGATCCATCCAAACAA	1200	
1201	ATTGTACAAATTGATGCTGAGAGATTGGAAATGTAAATTCAGGAAACAGAACTAAATCT	1260	
1201	ATTGTACAAATTGATGCTGAGAGATTGGAAATGTAAATTCAGGAAACAGAACTAAATCT	1260	
1261	TAA	1263	
1261	TAA	1263	

6935 standard; cDNA; 1263 BP.

AC ADB16935;
XX
DT 20-NOV-2003 (first entry)
XX
DE Gorilla DYX1L cDNA sequence.
XX
KW gene; ss; gorilla; DYX1L; dyslexia; neurological disorder;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Gorilla gorilla.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..1263
XX FT /*tag= a
XX FT /product= "DYX1L protein"
XX
XX WO2003068814-A1.
XX
XX 21-AUG-2003.
XX
XX 12-FEB-2003; 2003WO-FI000110.
XX
XX 12-FEB-2002; 2002US-0355782P.
XX
XX (LICN) LICENTIA LTD.
XX
XX Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
XX WPI; 2003-646482/61.
XX P-PSDB; ADB16936.
XX
XX New isolated, purified DYX1L nucleic acid for studying brain processes,
XX e.g. reading, phonological processing, rapid naming or verbal short-term
XX memory, or for diagnosing dyslexia or assessing the predisposition to
XX dyslexia.
XX
XX Claim 29; Page 119-121; 135pp; English.
XX
XX This invention relates to a novel isolated human gene DYX1L that is
XX functionally related to dyslexia, more particularly it describes single
XX nucleotide polymorphisms thought to predispose an individual in to
XX developing dyslexia. This is a neurological disorder with a genetic basis
XX (DYX1L has been isolated to chromosome 15q21), which manifests itself as
XX a specific reading disability. Specifically, DYX1L is can be useful in
XX study of brain processes such as reading, phonological processing, rapid
XX naming and verbal short-term memory. Accordingly, the present invention
XX describes methods and materials for analysing allelic variations in the
XX DYX1L gene, and also provides DYX1L as an antigen for the production of
XX antibodies used in the diagnosis of dyslexia. This polynucleotide
XX sequence is the gorilla DYX1L cDNA homologous to the human DYX1L gene of
XX the invention.
XX
SQ Sequence 1263 BP; 491 A; 212 C; 275 G; 285 T; 0 U; 0 Other;

Query Match 98.9%; Score 1248.6; DB 9; Length 1263;
Best Local Similarity 99.3%; Pred. No. 1.7e-260;
Matches 1254; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGGCGTCTTCTGTCT 60
DB 1 ATGCCCCCTTCAGTTAGCGATTACAGCTGGCAGCAGCAGAGACTGTGGTCTTCTGTCT 60

QY 61 CTGCCCTTCAAGGGGTGTGGTGTGAGACACGCGAGCTGTTCGACGGAAAACTATCTG 120
DB 61 CTGCCCTTCAAGGGGTGTGGTGTGAGACACGCGAGCTGTTCGACGGAAAACTATCTG 120

QY 121 AAGGTCAACTTCTCCATTTTATTTTGGAGCACTTTTATGTCTCCCATAGAGATGAG 180
DB 121 AAGGTCAACTTCTCCATTTTATTTTGGAGCACTTTTATGTCTCCCATAGAGATGAG 180

QY 181 AGCAGCAAGCAAGATTGGGAATGACACCACTTGTCTTACCTTTGTATAAAAAAGAGCG 240
DB 181 AGCAGCAAGCAAGATTGGGAATGACACCACTTGTCTTACCTTTGTATAAAAAAGAGCG 240

RESULT 6

ADB16937

ID ADB16937 standard; cDNA; 1263 BP.

XX ADB16937;

XX 20-NOV-2003 (first entry)

XX Orangutan DYXC1 cDNA sequence.

XX gene; ss; orangutan; DYXC1; dyslexia; neurological disorder;

XX reading disability; phonological processing; rapid naming;

XX verbal short-term memory.

XX OS Pongo pygmaeus.

XX Key Location/Qualifiers

XX 1..1263

XX /*tag= a

XX /product= "DYXC1 protein"

XX PN WO2003068814-A1.

XX PD 21-AUG-2003.

XX PF 12-FEB-2003; 2003WO-FI000110.

XX PR 12-FEB-2002; 2002US-0355782P.

XX XX (LICN) LICENTIA LTD.

XX Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;

XX WPI; 2003-646482/61.

XX DR P-PSDB; ADB16938.

New isolated, purified DYXC1 nucleic acid for studying brain processes, e.g. reading, phonological processing, rapid naming or verbal short-term memory, or for diagnosing dyslexia or assessing the predisposition to dyslexia.

PS Claim 29; Page 123-124; 135pp; English.

XX This invention relates to a novel isolated human gene DYXC1 that is functionally related to dyslexia, more particularly it describes single nucleotide polymorphisms thought to predispose an individual in to developing dyslexia. This is a neurological disorder with a genetic basis (DYXC1 has been isolated to chromosome 15q21), which manifests itself as a specific reading disability. Specifically, DYXC1 is can be useful in study of brain processes such as reading, phonological processing, rapid naming and verbal short-term memory. Accordingly, the present invention describes methods and materials for analysing allelic variations in the DYXC1 gene, and also provides DYXC1 as an antigen for the production of antibodies used in the diagnosis of dyslexia. This polynucleotide sequence is the orangutan DYXC1 cDNA homologous to the human DYXC1 gene of the invention.

XX Sequence 1263 BP; 493 A; 218 C; 272 G; 280 T; 0 U; 0 Other;

Query Match

Best Local Similarity 98.5%; Score 1243.8; DB 9; Length 1263;

Matches 1251; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy	1	ATGCTCTTCAGGTTAGCGATTACAGCTGGCGAGCAGACGAGACGTGGCGTCTTTCTGCT	60
Db	1	ATGCCCCCTCAGGTTAGCGATTACAGCTGGCGAGCAGACGAGACGTGGCGTCTTTCTGCT	60
Qy	61	CTGCCCTCAAGGGCGTGTGGCTCAGAGACAGGACGTGTCTGCACGGAAAACTATCTG	120
Db	61	CTGCCCTCAAGGGCGTGTGGCTCAGAGACAGGACGTGTCTGCACGGAAAACTATCTG	120

Qy	121	AAGTCAACTTTCCTCCATTTTATTTTGGAGCATTTCTTTATGCTCCCATAGACGATGAG	180
Db	121	AAGTCAACTTTCCTCCATTTTATTTTGGAGCATTTCTTTATGCTCCCATAGACGATGAG	180
Qy	181	AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTCACTTGATATAAAAAAGAGCG	240
Db	181	AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTCACTTGATATAAAAAAGAGCG	240
Qy	241	GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAAGAGATGATGCAAGAAATTAGA	300
Db	241	GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAAGAGATGATGCAAGAAATTAGA	300
Qy	301	GAAAAATCTATTTTACAGCACAAGAGAGACCAAGAGAGCTACAGAGCAAAAGCTGCA	360
Db	301	GAAAAATCTATTTTACAGCACAAGAGAGACCAAGAGAGCTACAGAGCAAAAGCTGCA	360
Qy	361	GCAAGCGGGAAGATCAAAAAATACGCACTAAGTGTGATGATGAAGATTGAAGAAAGAG	420
Db	361	GCAAGCGGGAAGATCAAAAAATATGCACTAAGTGTGATGATGAAGATTGAAGAAAGAG	420
Qy	421	AGGAAAAAATAGAGATATGAAAGAAATGAAACGGATATAAGCCACTAAAGCATTTGAA	480
Db	421	AGGAAAAAATAGAGATATGAAAGAAATGAAACGGATATAAGCCACTAAAGCATTTGAA	480
Qy	481	GCCTGGAAGAAATATCAAGAAAAAGCTGAGAGCAAAAAAATTCAGAGAGAGAGAA	540
Db	481	GCCTGGAAGAAATATCAAGAAAAAGCTGAGAGCAAAAAAATTCAGAGAGAGAGAA	540
Qy	541	TTATGTCAAAAAAGAAAGCAAAATTAAGAGAGAGAAAAAATAAAATATAAGAGTCTT	600
Db	541	TTATGTCAAAAAAGAAAGCAAAATTAAGAGAGAGAAAAAATAAAATATAAGAGTCTT	600
Qy	601	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAAATTCAGAAAAATATAT	660
Db	601	ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAAATTCAGAAAAATATAT	660
Qy	661	ACTGAGAAAGTTAAAGGAGACAGTATTCCTGCTCTGCTGCTCTGTTGGCAGTATTAATATC	720
Db	661	ACTGAGAAAGTTAAAGGAGACAGTATTCCTGCTCTGCTGCTCTGTTGGCAGTATTAATATC	720
Qy	721	AACCTTACCCCTCGAGTATTCCTCAACAGCTCTCGTGAATCAACAGTACAGAGAGAGAG	780
Db	721	AACCTTACCCCTCGAGTATTCCTCAACAGCTCTCGTGAATCAACAGTACAGAGAGAGAG	780
Qy	781	GAGTGGCTACACAAACAGCTGAGGCGACGAGACCAATGAATACTGACATAGCTGAACCTT	840
Db	781	GAGTGGCTACACAAACAGCTGAGGCGACGAGACCAATGAATACTGACATAGCTGAACCTT	840
Qy	841	TGCGATTTAAAG	900
Db	841	TGCGATTTAAAG	900
Qy	901	TTTCAACGGAAAACTATTTCGAGCTATCAATGCATATAATTTAGCCATAGACTAAAT	960
Db	901	TTTCAACGGAAAACTATTTCGAGCTATCAATGCATATAATTTAGCCATAGACTAAAT	960
Qy	961	AATAAGATGCCATATTGTTATTTGAACCGGGCTGCTTGCACCTAAATACTAAAAAATTA	1020
Db	961	AATAAGATGCCATATTGTTATTTGAACCGGGCTGCTTGCACCTAAATACTAAAAAATTA	1020
Qy	1021	CACAGGCTATTGAAGATTCTTTAAGGCACTGGAATTTATTTGATGCCACCTGTTACAGAC	1080
Db	1021	CACAGGCTATTGAAGATTCTTTAAGGCACTGGAATTTATTTGATGCCACCTGTTACAGAC	1080
Qy	1081	AATGCTAATGCAAGAAATGAAGGACATGTACGATGGAACAGCATTTCTGCACTAGAA	1140
Db	1081	AATGCTAATGCAAGAAATGAAGGACATGTACGATGGAACAGCATTTCTGCACTAGAA	1140
Qy	1141	TTGTATGTAGAGGCTCTACAGGATTATGAAGCGGCCTTTAAGATTGATTCATCAACAAA	1200
Db	1141	TTGTATGTAGAGGCTCTACAGGATTATGAAGCGGCCTTTAAGATTGATTCATCAACAAA	1200
Qy	1201	ATTGTACAAATTGATGTGAGAGAGATTTCGGGAATGTAATTCAGGAAACAGAACTAAATCT	1260

Db 1201 ATGTACAAATTGATGCTGAGAGATTTCGGAAATGTAATTCAGGAACAGAACTAAAATCT 1260
Qy 1261 TAA 1263
Db 1261 TAA 1263
RESULT 7
ADC30210
ID ADC30210 standard; cDNA; 1641 BP.
XX
XX ADC30210;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human novel cDNA sequence, SEQ ID NO:292.
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 15q21.3; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003029271-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030474.
XX
XX 24-SEP-2001; 2001US-0324631P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI; 2003-371981/35.
DR P-PSDB; ADC31181.
DR
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
XX Claim 1; SEQ ID NO 292; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30899) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of

CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1641 BP; 614 A; 272 C; 331 G; 424 T; 0 U; 0 Other;
Query Match 83.1%; Score 1049; DB 10; Length 1641;
Best Local Similarity 100.0%; Pred. No. 2.9e-217;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACAGAAAGACTGGGCTCTTCTCTCT 60
Db 104 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACAGAAAGACTGGGCTCTTCTCTCT 163
Qy 61 CTGCCCCCTCAAAGCGGTGCGTGCAGACACCGACGCTGTTCTCACGGAATACTACTG 120
Db 164 CTGCCCCCTCAAAGCGGTGCGTGCAGACACCGACGCTGTTCTCACGGAATACTACTG 223
Qy 121 AAGTCAACTTTCTCTCCATTTTATTGAGCGATTTCTTATGCTCCCATAGCATGAG 180
Db 224 AAGTCAACTTTCTCTCCATTTTATTGAGCGATTTCTTATGCTCCCATAGCATGAG 283
Qy 181 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGTTTCACTTTGTATAAAAAGAGCG 240
Db 284 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGTTTCACTTTGTATAAAAAGAGCG 343
Qy 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAATTAGA 300
Db 344 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAATTAGA 403
Qy 301 GAAAAATCTATTTTACAAGCAAGAGAGAGCAAAAGAGCTACAGAAGCAAAAGCTGCA 360
Db 404 GAAAAATCTATTTTACAAGCAAGAGAGAGCAAAAGAGCTACAGAAGCAAAAGCTGCA 463
Qy 361 GCAAGCGGGAAGATCAAAAATCGCACAATGATGATGATGATGATGATGATGATGATGAT 420
Db 464 GCAAGCGGGAAGATCAAAAATCGCACAATGATGATGATGATGATGATGATGATGATGAT 523
Qy 421 AGGAAAAAATAGAAAGATATGAAAGAAAATGAAACGGATATAAAGCCATTAAGCAATTCGAA 480
Db 524 AGGAAAAAATAGAAAGATATGAAAGAAAATGAAACGGATATAAAGCCATTAAGCAATTCGAA 583
Qy 481 GCCTGGAAGAATATCAAGAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAGAA 540
Db 584 GCCTGGAAGAATATCAAGAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAGAA 643
Qy 541 TTATGTCAAAAGAAAAGCAAAATTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 644 TTATGTCAAAAGAAAAGCAAAATTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 703
Qy 601 ACTAGAAAATTTGGCATCTAGAAAATCTTCTGCTCAAAAGGAGAGAAATTCAGAAAATATATTT 660
Db 704 ACTAGAAAATTTGGCATCTAGAAAATCTTCTGCTCAAAAGGAGAGAAATTCAGAAAATATATTT 763
Qy 661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCTCTGCTCTGTCAGTATTAAATATC 720
Db 764 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCTCTGCTCTGTCAGTATTAAATATC 823
Qy 721 AACTTTTCCCTCCAGTATTCCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 780
Db 824 AACTTTTCCCTCCAGTATTCCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 883
Qy 781 GAGTGGCTACACAAACAGCTGAGGCGACGAGCAATGAATATCTGACATAGCTGAACCT 840

Db 884 GAGTGGCTACACAAACAGCTGAGCCACGAAGAGCAATGAATCTGACATAGCTGAACCTT 943
Qy 841 TGGGATTTAAAGAAAGAAAGAACCCAGAAATGGTTGAAGGATAAAGGAAACAAATTTG 900
Db 944 TGGGATTTAAAGAAAGAAAGAAACCCAGAAATGGTTGAAGGATAAAGGAAACAAATTTG 1003
Qy 901 TTGCAACGGGAACATAATTTGGCAGCTATCATGCAATATATTTAGCCATAGACTAAAT 960
Db 1004 TTGCAACGGGAACATAATTTGGCAGCTATCATGCAATATATTTAGCCATAGACTAAAT 1063
Qy 961 AATAAGATGCCACTATTTGATTTGAACCGGGCTGCTTGCACCTAAAACTAAAAAATTTA 1020
Db 1064 AATAAGATGCCACTATTTGATTTGAACCGGGCTGCTTGCACCTAAAACTAAAAAATTTA 1123
Qy 1021 CACAAGGCTATTGAAGATCTTCTTAAGC 1049
Db 1124 CACAAGGCTATTGAAGATCTTCTTAAGC 1152

RESULT 8

ADMO1890
ID ADM01890 standard; cDNA; 1559 BP.
XX AC
XX AC
XX ADM01890;
XX
DT 20-MAY-2004 (first entry)
XX DE Human cDNA of the invention SEQ ID NO:575.
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX Homo sapiens.
XX EP1347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JP-00137785.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI; 2003-723558/69.
XX P-FSDB; ADM04333.

XX New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 575; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX cDNA sequence of the invention.

XX Sequence 1559 BP; 592 A; 252 C; 309 G; 406 T; 0 U; 0 Other;

XX Query Match 82.9%; Score 1047.4; DB 11; Length 1559;
XX Best Local Similarity 99.9%; Pred. No. 6.4e-217;
XX Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 60
Db 43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 102
Qy 61 CTGCCCCCTCAAAGCGGTGTGGTTCAGAGACACGACGCTGTTCTGCAACGAAAACTATCTG 120
Db 103 CTGCCCCCTCAAAGCGGTGTGGTTCAGAGACACGACGCTGTTCTGCAACGAAAACTATCTG 162
Qy 121 AAGTCAACTTTTCTCCATTTTATTTGAGGCACTTTCTTTATGCTCCCATAGACGATGAG 180
Db 163 AAGTCAACTTTTCTCCATTTTATTTGAGGCACTTTCTTTATGCTCCCATAGACGATGAG 222
Qy 181 AGCAGCAAGCAAGATTTGGGAATGACACCACTTCTTACCTTGTATATAAAGAAAGCG 240
Db 223 AGCAGCAAGCAAGATTTGGGAATGACACCACTTCTTACCTTGTATATAAAGAAAGCG 282
Qy 241 GCCATGTGGGAGACCTTTCTGTACGGGTGTGCAAAAGAGATGATGCAAGGAATAGA 300
Db 283 GCCATGTGGGAGACCTTTCTGTACGGGTGTGCAAAAGAGATGATGCAAGGAATAGA 342
Qy 301 GAAAAATCTATTTTACAAGCAAAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCA 360
Db 343 GAAAAATCTATTTTACAAGCAAAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCA 402
Qy 361 GCAAAGCGGAGAGATCAAAAATACGCATAAGTGTCTGATGAAGATGAAGAGAGAG 420
Db 403 GCAAAGCGGAGAGATCAAAAATACGCATAAGTGTCTGATGAAGATGAAGAGAGAG 462
Qy 421 AGGAAAAAATAGAAAGATATGAAAGAAAAATGAAAGGCAATGAAAGCAATTTGAA 480
Db 463 AGGAAAAAATAGAAAGATATGAAAGAAAAATGAAAGGCAATGAAAGCAATTTGAA 522
Qy 481 GCCTGGAAAGAAATATCAAGAAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 540
Db 523 GCCTGGAAAGAAATATCAAGAAAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAAA 582
Qy 541 TTATGTCAAAAAGAAAAAGCAAAATTAAGAGAGAGAAAAAATAAATAATATAAGAGCTT 600
Db 583 TTATGTCAAAAAGAAAAAGCAAAATTAAGAGAGAGAAAAAATAAATAATATAAGAGCTT 642
Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTGTCTCAAAAGGAGAAATTCAGAAAAATATTT 660
Db 643 ACTAGAAATTTGGCATCTAGAAATCTTGTCTCAAAAGGAGAAATTCAGAAAAATATTT 702
Qy 661 ACTGAGAAGTTAAAGGAGAGAGATTTCTGCTCTCTCTCTGTTGGCAGTATTAAAAATC 720
Db 703 ACTGAGAAGTTAAAGGAGAGAGATTTCTGCTCTCTCTCTGTTGGCAGTATTAAAAATC 762
Qy 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTGTGAATCACAAGTAGCAGAGAGAG 780
Db 763 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTGTGAATCACAAGTAGCAGAGAGAG 822
Qy 781 GAGTGGCTACACAAACAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAACCTT 840
Db 823 GAGTGGCTACACAAACAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAACCTT 882
Qy 841 TGGCATTTAAAGAAAGAAAAAGAACCCAGAAATGGTTGAAGGATAAAGGAAAAACAATTTG 900
Db 883 TGGCATTTAAAGAAAGAAAAAGAACCCAGAAATGGTTGAAGGATAAAGGAAAAACAATTTG 942
Qy 901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAGACTAAAT 960
Db 943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAGACTAAAT 1002
Qy 961 AATAAGATGCCACTATTTGATTTGAACCGGGCTGCTTGCACCTAAAACTAAAAAATTTA 1020
Db 1003 AATAAGATGCCACTATTTGATTTGAACCGGGCTGCTTGCACCTAAAACTAAAAAATTTA 1062
Qy 1021 CACAAGGCTATTGAAGATTTCTTCTTAAGGC 1049
Db 1063 CACAAGGCTATTGAAGATTTCTTCTTAAGGC 1091

```
RESULT 9
ADB16924
ID ADB16924 standard; cDNA; 1697 BP.
XX
XX AC ADB16924;
XX
XX 20-NOV-2003 (first entry)
XX
XX cDNA sequence of the murine DYXC1 mRNA.
XX
XX gene; ss; mouse; DYXC1; dyslexia; neurological disorder;
XX reading disability; phonological processing; rapid naming;
XX verbal short-term memory; murine.
XX
XX OS Mus musculus.
XX
XX FH Location/Qualifiers
XX FT 48..1310
XX FT /*tag= a
XX FT /product= "DYXC1 protein"
XX
XX WO2003068814-A1.
XX
XX 21-AUG-2003.
XX
XX 12-FEB-2003; 2003WO-FI000110.
XX
XX 12-FEB-2002; 2002US-0355782P.
XX
XX (LICN ) LICENTIA LTD.
XX
XX Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
XX WPI; 2003-646482/61.
XX DR P-PSDB; ADB16925.
XX
XX New isolated, purified DYXC1 nucleic acid for studying brain processes,
XX e.g. reading, phonological processing, rapid naming or verbal short-term
XX memory, or for diagnosing dyslexia or assessing the predisposition to
XX dyslexia.
XX
XX Disclosure; Page 51-53; 135pp; English.
XX
XX This invention relates to a novel isolated human gene DYXC1 that is
XX functionally related to dyslexia, more particularly it describes single
XX nucleotide polymorphisms thought to predispose an individual in to
XX developing dyslexia. This is a neurological disorder with a genetic basis
XX (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
XX a specific reading disability. Specifically, DYXC1 is can be useful in
XX study of brain processes such as reading, phonological processing, rapid
XX naming and verbal short-term memory. Accordingly, the present invention
XX describes methods and materials for analysing allelic variations in the
XX DYXC1 gene, and also provides DYXC1 as an antigen for the production of
XX antibodies used in the diagnosis of dyslexia. This polynucleotide is the
XX cDNA sequence of the murine DYXC1 mRNA of the invention.
XX
XX Sequence 1697 BP; 534 A; 360 C; 441 G; 362 T; 0 U; 0 Other;
XX
XX Query Match 66.6%; Score 840.8; DB 9; Length 1697;
XX Best Local Similarity 80.3%; Pred. No. 3.5e-172;
XX Matches 1012; Conservative 0; Mismatches 242; Indels 6; Gaps 2;
XX
XX 1 ATGCCTCTTACAGTTAGCGATTACAGCTGGCAGACGAGCACTGGGCTTTCTGTCT 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 48 ATGCAGTGCAGTGCAGGAATTCAGCTGGCAGCAGACGCCGGCCAGATCTCTCTGCG 107
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 CTGCCCCCTCAAGCGCTGCTGCAGACACGACGCTGTTCTGCACGGAAACTATCTG 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 108 CTGCTCTGCGGGGGCTGCTGGCGCGATCTGACGATTTCTGTGGGGAAGTACCTG 167
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 AAGGTCAACTTTCCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGCAGTAG 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 168 AAGGTTAACTTTCCTCCATTTTATTTGAGCTGTTTCTCTATGCTCCCATAGATGGG 227
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |

181 AGCAGCAAAAGCAAGATTGGGAATGACACCAATTTGTCTTACCTTGTATATAAAGAAGCG 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
228 AAGAGCAAAAGCCAAGATTGGAAATGACACCAATTTCTTTTTCATTTGATATAAAGAGGCCA 287
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 GCCATGTGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAAGATTAGA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
288 GTTCTGTGGGATAGCCTTTCTGTGCGGGTGTGATAAAGAGATGATGACAGAAATAGA 347
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 GAAAAATCTATTTTACAGCACAAGAGAGCAAAAGAAAGCTACAGAGCAAAAGCTGCA 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
348 GAAAAATCTATCTTTCAGCACAAGAGAAAGCAAAAGAGGCCACAGAAAGCAAAAGCTGT 407
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 GCAAGCGGGAAGATCAAAAATACGCACAAAGTGTATGATGAAGATTTGAAGAAGAAGAG 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
408 GCCAAGCGAGAGACCGAGATACGCACCTAGGCGAGATGATGAAGATTTGAAGAAGAAGAG 467
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 AGGAAAAAATAGAGATATGAAAGAAATGAAACGGATATAAAGCCATTAAGCATTTGAA 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
468 AGGAAAAAATCTCGAAGATCTGAAAGAAATGAAACGGAAAGGCACTAGCGAATTAGAA 527
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 GCCTGGAAGAATATCAAGAAAGCTGAGAGCAAAAGAAAAAATTTAGAGAGAGAGAAA 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
528 GCCTGGAAGAATGTCAAAAGAGAAAGCTGACCGCAAAAGAAAGAGTCCAGAGGAAGAGAA - 586
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGGAAGAAAAAATAAATAATTAAGAGTCTT 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
587 --ACCGCTCGAGGAAAGCAAGCT---GAAGAGACCAAGCTCTAATAAATCTCGGGTTTG 641
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
642 CCCGGAAGGCCCCACCCTCCCTCCCAAGAGGGAGGAATTCGGAAAAACATATTT 701
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTGTCAGTATTAATATC 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
702 CCTGAGAAGTTAAAGGAAGACAGAGTCCCTCGCTCGCTCGCTGCGAGTATTCAAATC 761
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
721 AACTTTACCCCTCGAGTATTTCCCAAGAGCTCTTCGTGAATCAACAAGTAGCAGAGAGAG 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
762 AGCTTTACCCCTCGAGTATTTCCCAAGAGCTCTTCGGGAATCCCAAGTCGCGAAGAGAG 821
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
781 GAGTGGCTACACAAAGCTGAGGCAAGAGCAATGAATACTGACATAGCTGAACCTT 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
822 GAGTGGCTGCATAAACAAAGCAGAGACACGAGAGCCATGAGCACTGACCTCTCTGAGTTC 881
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
841 TGGATTTAAAGAGAGAAAGAACCCAGATGTTGAGATTAAGAGTAAGAGAAACAAATG 900
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
882 TTTGACTTAAAGAGAGAGAGGAATCCAGACTGGTTGAAAGACAAAGGGAAACAAATG 941
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
901 TTTGCAACGGAAACTATTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
942 TTTGCAACAGAAACTATTTGGCAGCTGATGATGATATATTTAGCCATACGACTGAC 1001
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
961 AATAAGATGCCACTATTTGTAATTTGAACCCGGCTGCTTGCACCTTAAACCTAAAAA 1020
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1002 TGTAAAGATCCCATTTATTTGATTTGAATCGGGCTGCTTGCACCTCAAAATTA 1061
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1021 CACAGGCTATTGAGATTTCTTAAGGCACTGGAATTTATGATGCCACCTGTTTACAGAC 1080
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1062 CACAAGGCCATCGAGGACTCTTCTTAAGGCACTAGAGTATTTGACACCACTGTTGCTGAC 1121
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1081 AATGCTAATCAAGAAATGAAGGCACATGTAGCAGCTGGAACAGCATTTCTGTCAACTAGAA 1140
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1122 AATGCCAATGCAAGAAATGAAGGCACATGTAGCAGGAGGACAGGCTTCTGTCACTAGAA 1181
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1141 TTGTATGTAGAGGCCCTACAGGATTTAAGAGCGGCCTTAAGATTTGATCCATCCAA 1200
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1182 TTGTATGTAGAGGCCCTTGAAGATTTAAGAGCTGCACTTAAGATTTGACCCAGCCACACA 1241
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1201 ATTGTACAAATTTGATGTGAGAGATTTGGAATTTGATTTCAAGGAAACAGACTAAATCT 1260
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
1242 GTTGTACAGAACGATGAGAGAGATTTGGAATTTAATTTCAAGGGACGGCACTGAAGTCT 1301
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```


RESULT 10
 ADC32116
 ID ADC32116 standard; cDNA; 608 BP.
 XX
 AC ADC32116;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel cDNA contig sequence, SEQ ID NO:2198.
 XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; anti-anemic; anticoagulant; thrombolytic; vulnerary;
 KW anticulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 15q21.3; 88.
 XX
 OS Homo sapiens.
 XX
 XX WO2003029271-A2.
 PN
 XX
 PD 10-APR-2003.
 XX
 XX 24-SEP-2002; 2002WO-US030474.
 PF
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QH, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 XX WPI; 2003-371981/35.
 DR P-PSDB; ADC32883.
 XX
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 XX Example 2; SEQ ID NO 2198; 1185pp; English.
 PS
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,

CC and as food supplements. The present sequence represents a human contig
 CC sequence used in an example of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 608 BP; 230 A; 103 C; 135 G; 140 T; 0 U; 0 Other;

Query Match 41.5%; Score 524; DB 10; Length 608;
 Best Local Similarity 100.0%; Pred. No. 1.1e-103;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 TTCAGAGAGAGAGAAATTATGTCARAAAGAAAGCAAAATTAAGAAGAGAGAAAAAAA 583
 DB 1 TTCAGAGAGAGAGAAATTATGTCARAAAGAAAGCAAAATTAAGAAGAGAGAAAAAAA 60
 QY 584 TAAATATAAGAGCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 643
 DB 61 TAAATATAAGAGCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 120
 QY 644 ATTCAGAAATATATTTACTGAGAAAGTTAAAGGAGACAGTATTCTCTCTCTCTCTG 703
 DB 121 ATTCAGAAATATATTTACTGAGAAAGTTAAAGGAGACAGTATTCTCTCTCTCTCTG 180
 QY 704 TTGGCAGTATTAAATCAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCAC 763
 DB 181 TTGGCAGTATTAAATCAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCAC 240
 QY 764 AAGTAGCAGAGAGAGAGGAGTGCTACACAAACAGCTGAGGCAGAGAGAGCAATGAATA 823
 DB 241 AAGTAGCAGAGAGAGAGGAGTGCTACACAAACAGCTGAGGCAGAGAGAGCAATGAATA 300
 QY 824 CTGCATAGCTGAACTTTTGGCATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 883
 DB 301 CTGCATAGCTGAACTTTTGGCATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 884 ATAAAGGAG 943
 DB 361 ATAAAGGAG 420
 QY 944 TAGCCATAAG 1003
 DB 421 TAGCCATAAG 480
 QY 1004 TAAAACTAAAAAACTTACACAGGCTATTGGAAGATTCTTCTTAAG 1047
 DB 481 TAAAACTAAAAAACTTACACAGGCTATTGGAAGATTCTTCTTAAG 524

RESULT 11
 ACH35463
 ID ACH35463 standard; cDNA; 488 BP.

XX ACH35463;
 AC

DT 13-OCT-2003 (first entry)

DE Human endothelial cell cDNA #3596.

KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

XX US2003073623-A1.

PD 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

PR 30-JUL-2001; 2001US-00918995.

PA (DRMA/) DRMANAC R T.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:40:13 ; Search time 4597 Seconds
(without alignments)

10457.945 Million cell updates/sec

Title: US-10-681-199-1

Perfect score: 1263
Sequence: 1 atgctctctcaggtagcgca.....gaacagaactaaatcttaa 1263

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1034.8	81.9	1600	3	BC017392 Homo sapi
2	812.6	64.3	844	6	CD107587 AGENCOURT
3	676.4	53.6	933	6	CD358543 AGENCOURT
4	667.8	52.9	793	7	CO738958 SILE04c21
5	567.4	44.9	792	4	BG771796 602720472
6	553	43.8	788	2	BES64350 601343161
7	519	41.1	683	5	BP460416 BP460416
8	513.8	40.7	1190	3	AK005832 Mus muscu
9	510.4	40.4	1131	5	BQ217312 AGENCOURT
10	508	40.2	917	6	BY705981 BY705981
11	488.6	38.7	545	2	BF216970 601884034
12	474.4	37.6	535	7	CN478982 UI-CF-FNO
13	449	35.6	796	7	CK603982 AGENCOURT
14	434.2	34.4	468	6	CB297042 12B22058
15	427	33.8	1183	2	BF207765 601861861
16	422	33.4	462	7	CN429788 170006000
17	414.4	32.8	417	1	AI024221 cv81e02.s
18	409.6	32.4	988	4	BI517373 603041624
19	408	32.3	559	5	BU607404 UI-CF-FNO
20	405	32.1	823	5	BU567908 AGENCOURT
21	395.6	31.3	1153	3	AK015345 Mus muscu
22	389.2	30.8	701	4	BG482634 602502783
23	385.8	30.5	651	5	BU852580 AGENCOURT
24	376.8	29.8	844	7	CK598370 AGENCOURT

c	25	374.4	29.6	772	6	CA422987	UI-H-FLO-
	26	371.8	29.4	665	4	BG540324	BG540324 602568825
	27	364	28.8	573	2	BE972748	BE972748 601652170
	28	363.2	28.8	775	4	BG206612	BG206612 RST26063
	29	350.8	27.8	477	6	CB545994	CB545994 AMGNNUC-M
	30	344.4	27.3	744	4	BG192162	BG192162 RST11269
	31	341	27.0	752	4	BG183140	BG183140 RST2158 A
c	32	325	25.7	325	1	AI783611	AI783611 t299601.x
	33	308.4	24.4	375	2	AW481240	AW481240 35107 MAR
	34	301.8	23.9	687	4	BG242087	BG242087 602354631
	35	298	23.6	764	5	BU307742	BU307742 603539279
	36	296.2	23.5	767	5	BU307742	BU307742 603539279
	37	295	23.4	855	2	BF248143	BF248143 601859338
c	38	293.2	23.2	727	4	BU056709	BU056709 BJ056709
	39	291.8	23.1	478	2	BB615485	BB615485 BB615485
	40	291.8	23.1	478	3	AK076862	AK076862 Mus muscu
	41	288.4	22.8	543	7	CK137348	CK137348 mak06b02.
	42	283.8	22.5	470	1	AV266346	AV266346 AV266346
	43	273.2	21.6	375	2	AW415915	AW415915 50508 MAR
	44	270.8	21.4	617	5	BU267200	BU267200 603819026
	45	269.4	21.3	370	1	AA967396	AA967396 ua05c07.r

ALIGNMENTS

RESULT 1
BC017392
LOCUS BC017392 1600 bp mRNA linear HTC 23-AUG-2002
DEFINITION Homo sapiens, Similar to RIKEN cDNA 1700010124 gene, clone IMAGE:4081622, mRNA.
ACCESSION BC017392
VERSION BC017392.1 GI:19263480
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
<http://www.sysatembiology.org>
contact: amadaneysystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 32 Row: k Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction
This clone has the following problem: frame shifted.

FEATURES
source Location/Qualifiers
1..1600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4081622"
/tissue_type="Bladder, carcinoma"
/clone_lib="NIH_MGC_53"
/lab_host="DH10B"

/note="Vector: pDNR-LIB"			
ORIGIN			
Query Match		81.9%; Score 1034.8; DB 3; Length 1600;	
Best Local Similarity		99.7%; Pred. No. 9.3e-215;	
Matches 1047; Conservative		0; Mismatches 2; Indels 1; Gaps 1;	
Qy	1	ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACGACGAGAGCTGCGGTCTTTCTCTCT	60
Db	53	ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGACGACGAGAGCTGCGGTCTTTCTCTCT	112
Qy	61	CTGCCCTCAAAGCGGTGTGCGTCAGAGACACGACGCTGTTCTGCACGGAAAACTATCTG	120
Db	113	CTGCCCTCAAAGCGGTGTGCGTCAGAGACACGACGCTGTTCTGCACGGAAAACTATCTG	172
Qy	121	AAGTCAACTTCTCCATTTTATTTTATTTAGGCAATTTCTTATGTCTCCATAGAGATGAG	180
Db	173	AAGTCAACTTCTCCATTTTATTTTATTTAGGCAATTTCTTATGTCTCCATAGAGATGAG	232
Qy	181	AGCAGCAAGCAAGATTTGGGAATGACACCAATGCTCTTACCTTGTTATAAAAAAGAGCG	240
Db	233	AGCAGCAAGCAAGATTTGGGAATGACACCAATGCTCTTACCTTGTTATAAAAAAGAGCG	292
Qy	241	GCCATGTGGGAGACCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAAGATTAGA	300
Db	293	GCCATGTGGGAGACCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAAGATTAGA	352
Qy	301	GAATAATCTATTTTACAAGCACAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCA	360
Db	353	GAATAATCTATTTTACAAGCACAAGAGAGACAAAGAGCTACAGAGCAAAAGCTGCA	412
Qy	361	GCAAGCGGGAAGATCAAAATACGCACTAGTGTCTATGATGAAGATTTGAAGAAGAAGAG	420
Db	413	GCAAGCGGGAAGATCAAAATACGCACTAGTGTCTATGATGAAGATTTGAAGAAGAAGAG	472
Qy	421	AGG-AAAAAATAGAGATATGAAGAAAAATGAACGGATAAAGCCACTAAAGCATTTGGA	479
Db	473	AGGAAAAAATAGAGATATGAAGAAAAATGAACGGATAAAGCCACTAAAGCATTTGGA	532
Qy	480	AGCTGGAAGAATATCAAGAAAAAGCTGAGGAGCAAAAAAATAATTCAGAGAGAAGAA	539
Db	533	AGCTGGAAGAATATCAAGAAAAAGCTGAGGAGCAAAAAAATAATTCAGAGAGAAGAA	592
Qy	540	ATTATGTCAAAAGAAAGCAAAATTAAGAGAGGAAGAAAAAATAAATAATAGAGTCT	599
Db	593	ATTATGTCAAAAGAAAGCAAAATTAAGAGAGGAAGAAAAAATAAATAATAGAGTCT	652
Qy	600	TACTAGAAATTTGGCATCTAGAAATCTTTGCTCCAAAAGGGAGAAATTCAGAAAAATATT	659
Db	653	TACTAGAAATTTGGCATCTAGAAATCTTTGCTCCAAAAGGGAGAAATTCAGAAAAATATT	712
Qy	660	TACTAGAAATTTAAGGAAGACAGATATTCCTGCTCTGCTCTGTTGGCAGTATTAAT	719
Db	713	TACTAGAAATTTAAGGAAGACAGATATTCCTGCTCTGCTCTGTTGGCAGTATTAAT	772
Qy	720	CAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTCGAATTCACAGTAGCAGAGGA	779
Db	773	CAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTCGAATTCACAGTAGCAGAGGA	832
Qy	780	GGAGTGGCTACAAAAAAGCTGAGGACCAAGAGCAATGAATATCTGACATAGCTGAAT	839
Db	833	GGAGTGGCTACAAAAAAGCTGAGGACCAAGAGCAATGAATATCTGACATAGCTGAAT	892
Qy	840	TTGCGATTTAAAAAGAAAGAAAAAGAACCCAGAAATGGTTGAAGGATTAAGGAAACAAAT	899
Db	893	TTGCGATTTAAAAAGAAAGAAAAAGAACCCAGAAATGGTTGAAGGATTAAGGAAACAAAT	952
Qy	900	TTTTGCAACGAAACTATTTGGCAGCTATCAATGCTATATATTTAGCCATAAGCTAAA	959
Db	953	TTTTGCAACGAAACTATTTGGCAGCTATCAATGCTATATATTTAGCCATAAGCTAAA	1012
Qy	960	TAATAAGATGCCACTATTGTATTTTGAACCGGGCTGCTTGCACCCCTTAAACCTT	1019
Db	1013	TAATAAGATGCCACTATTGTATTTTGAACCGGGCTGCTTGCACCTTAAACCTT	1072
Qy	1020	ACACAGGCTATTGAAGATTCTTTCTAAGGC	1049
Db	1073	ACACAGGCTATTGAAGATTCTTTCTAAGGC	1102
RESULT 2			
LOCUS	CD107587	844 bp	mRNA linear EST 15-MAY-2003
DEFINITION	AGENCOURT 14021179 NIH_MGC_179 Homo sapiens cDNA clone		
	IMAGE:30368368 5', mRNA sequence.		
ACCESSION	CD107587		
VERSION	CD107587.1	GI:30760665	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 844)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: Dr. Michael Brownstein CDNA Library Preparation: Invitrogen Corp CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDAM431 row: a column: 17 High quality sequence stop: 634.		
FEATURES		Location/Qualifiers	
source		1..844	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:30368368"	
		/tissue_type="Pituitary"	
		/lab_host="DHI08-Ton A (T1 and T5 phage resistances)"	
		/clone_lib="NIH_MGC_179"	
		/note="Organ: brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."	
ORIGIN			
Query Match		64.3%; Score 812.6; DB 6; Length 844;	
Best Local Similarity		99.3%; Pred. No. 2.2e-166;	
Matches 837; Conservative		0; Mismatches 4; Indels 2; Gaps 2;	
Qy	18	CGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCTCTGCCCCCTCAAAGGCGT	77
Db	1	CGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCTCTGCCCCCTCAAAGGCGT	60
Qy	78	GTGCGTCAGAGACACGACGCGTGTTCGACGGAATACTATCTGAAGTCAACTTTCCTCC	137
Db	61	GTGCGTCAGAGACACGACGCGTGTTCGACGGAATACTATCTGAAGTCAACTTTCCTCC	120
Qy	138	ATTTTATTTTGGGCATTTCTTTATGCTCCCATAGCGATGAGAGCAAGCAAGAT	197
Db	121	ATTTTATTTTGGGCATTTCTTTATGCTCCCATAGCGATGAGAGCAAGCAAGAT	180
Qy	198	TGGGAATGACACCACTTGTCTTACCTTGTATATAAAAAAGAGCGCCATGTGGGAGACCT	257
Db	181	TGGGAATGACACCACTTGTCTTACCTTGTATATAAAAAAGAGCGCCATGTGGGAGACCT	240
Qy	258	TTCTGTGCGGGTCTTGACAAAGAGATGATGCAAAAGAAATTAGAGAAAAATCTATTACA	317
Db			

```
Db 241 TTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTAGAGAAAATCTATTTTACA 300
Qy 318 AGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGAGCAAAAGCGGGAAGATCA 377
Db 301 AGCAACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGAGCAAAAGCGGGAAGATCA 360
Qy 378 AAAATACCACTTAAGTGTCTATGATCAAGATTTGAAGAAGAGAGAGGAAAATAAGAGA 437
Db 361 AAAATACGCACCTAAGTGTCTATGATCAAGATTTGAAGAAGAGAGAGGAAAATAAGAGA 420
Qy 438 TATGAAAGAAAATGAACGGATAAAAGCACTAAAGCATTTGGAAGCTTGAAAGAAATATCA 497
Db 421 TATGAAAGAAAATGAACGGATAAAAGCACTAAAGCATTTGGAAGCTTGAAAGAAATATCA 480
Qy 498 AAGAAAAGCTGAGGAGC-AAAAAAAATTCAGAGAGAGAGAGAAATATATGTCAAAAAGAAA 556
Db 481 AAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAGAGAGAAATATATGTCAAAAAGAAA 540
Qy 557 AGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
Db 541 AGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 617 CTAGAAAATCTTGCTCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAATTAAGAG 676
Db 601 CTAGAAAATCTTGCTCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAATTAAGAG 660
Qy 677 AAGACAGATATCTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
Db 661 AAGACAGATATCTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 737 TATTTCCCAACAGCTTTGCTGATCACAAGTAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
Db 721 TATTTCCCAACAGCTTTGCTGATCACAAGTAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 797 AAGCTGAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 855
Db 781 AAGCTGAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 856 GAA 858
Db 841 GAA 843

RESULT 3
CD358543
LOCUS
DEFINITION AGNCOURT_14255831 NIH_MGC_180 Homo sapiens cDNA clone
IMAGE:30386203 5', mRNA sequence.
ACCESSION CD358543
VERSION CD358543.1 GI:31129978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@nci.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM454 row: h column: 20
High quality sequence start: 303
```

```
FEATURES
source
High quality sequence stop: 666.
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cdones="IMAGE:30386203"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site: 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."
```

ORIGIN

	Query Match	53.6%	Score 676.4	DB 6	Length 933
	Best Local Similarity	96.5%	Pred. No. 1e-136		
	Matches 723	Conservative 0	Mismatches 22	Indels 4	Gaps 3
Qy	159	TTATGCTCCCATAGACGATGAGAGCAGCAAAAGATTGGGAATGACACCATTTGCTT	218		
Db	33	TTATGCTCCCATAGACGATGAGAGCAGCAAAAGATTGGGAATGACACCATTTGCTT	92		
Qy	219	CACCTTGATATAAAAAA-GAAGCGGCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACA	277		
Db	93	CACCTTGATATAAAAAAACCAGCGGCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACA	152		
Qy	278	AAGAGATGATCAAGAAATTTAGAGAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAG	337		
Db	153	AAGAGATGATCAAGAAATTTAGAGAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAG	212		
Qy	338	AAGCTACAGAAAGCAAAAGCTGAGCAAAAGCGGGAAGATCAAAAATACGCATTAAGTGCA	397		
Db	213	AAGCTACAGAAAGCAAAAGCTGAGCAAAAGCGGGAAGATCAAAAATACGCATTAAGTGCA	272		
Qy	398	TGATGAAGATTGAAGAAG	456		
Db	273	TGATGAAGATTGAAGAAG	332		
Qy	457	ATAAAGCCACTTAAGCAATTTGGAAGCTGGAAGAAATATCAAGAAAGAGCTGAGGAGCA	516		
Db	333	ATAAAGCCACTTAAGCAATTTGGAAGCTGGAAGAAATATCAAGAAAGAGCTGAGGAGCA	392		
Qy	517	AAAAAATTCAG	576		
Db	393	AAAAAATTCAG	452		
Qy	577	AAAAAATTAAGAGATCTTCTAGAAAATTTGGCATCTAGAAAATCTTGGCTCAAAA	636		
Db	453	AAAAAATTAAGAGATCTTCTAGAAAATTTGGCATCTAGAAAATCTTGGCTCAAAA	512		
Qy	637	GGGGAATTCAGAAAATATATTTACTGAGAGTTTAAGGAGAGAGAGAGAGAGAGAGAGAG	696		
Db	513	GGGGAATTCAGAAAATATATTTACTGAGAGTTTAAGGAGAGAGAGAGAGAGAGAGAGAG	572		
Qy	697	CGCTCTGTGGCAGTATTAATAATCAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTGGT	756		
Db	573	CGCTCTGTGGCAGTATTAATAATCAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTGGT	632		
Qy	757	GAATCACAAGTAGCAG	816		
Db	633	GAATCACAAGTAGCAG	692		
Qy	817	ATGAATCTGATAGCTGAATCTTGGCATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	874		
Db	693	ATGAATCTGATAGCTGAATCTTGGCATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	752		
Qy	875	GGTTGAAGGATAAAGGAG	903		
Db	753	GGTTGAAGGATAAAGGAG	781		


```

RESULT 4
CO738958          793 bp  mRNA  linear  EST 29-JUL-2004
LOCUS             SILE04c21p08f1 squirrel embryo library 1 Spermophilus lateralis
DEFINITION        cDNA clone 21p08 5', mRNA sequence.
ACCESSION         CO738958
VERSION           CO738958.1 GI:50826228
KEYWORDS          EST.
SOURCE            Spermophilus lateralis (golden-mantled ground squirrel)
ORGANISM          Spermophilus lateralis
AUTHORS           Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
                  Rogers,J. and Cossins,A.R.
TITLE             Microarray analysis of transcriptional changes during hibernation
                  in the golden mantled ground squirrel, Spermophilus lateralis
JOURNAL           Unpublished (2004)
COMMENT           Contact: Andrew R. Cossins
                  Laboratory for Environmental Gene Regulation
                  University of Liverpool
                  School of Biological Sciences, The Biosciences Building, Crown
                  Street, Liverpool, United Kingdom, L69 7ZB
                  Tel: +44 (0)151-795-4510
                  Fax: +44 (0)151-795-4431
                  Email: cossins@liv.ac.uk
                  Vector has been trimmed from this EST.
                  Plate: 21 row: p column: 08
                  Seg primer: pflc T7 (5'-AATAGACTCACTATAGGG-3')
                  High quality sequence stop: 793.
FEATURES          Location/Qualifiers
                    1..793
                        /organism="Spermophilus lateralis"
                        /mol_type="mRNA"
                        /db_xref="taxon:76772"
                        /clone="21p08"
                        /sex="male and female"
                        /tissue_type="embryo"
                        /dev_stage="embryonic"
                        /lab_host="E.coli Electromax DH10B"
                        /clone_lib="squirrel embryo library 1"
                        /note="Vector: pFLC; Site 1: Sali GTCGAG; Site 2: BamHI
                        GCATCC; Normalized and subtracted cDNA library prepared
                        from embryos"
ORIGIN
Query Match      52.9%; Score 667.8; DB 7; Length 793;
Best Local Similarity 90.3%; Pred. No. 7.5e-135;
Matches 714; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 346 GAAGCAAAAGCTGAGCAAGCGGGAAGATCAAAAATACGCACTAAGTGTTCATGATGAAG 405
Db 2 GAAGCAAAAGCTACAGCAAGCGGAGAGATCAAAAATATACACTAAATGTTCATGATGCAG 61
Qy 406 ATTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Db 62 ATTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
Qy 466 ACTAAAGCATTTGGAAGCTCGAAAGATATCAAGAGAAAGCTGAGGAGCAAAAAAATT 525
Db 122 ACTAAGAGATTTGGAAGCTCGAAAGATGTCAGAGAAAGCTGAGAGACGAGAAAGATTT 181
Qy 526 CAGAGAGAGAGAGAAATATTGTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
Db 182 CAGAGAGAGAGAGAAATATTACGTCAACAGAGAAAGCAATTTGACGAGAGAGAGAGAGAG 241
Qy 586 AATATAGAGCTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCGAGAGAGAGAGAT 645
Db 242 AGACATAAATCTCATTTAAAAATTTCAACATCTTAAAAATCTTGCCACAAAAAGGAGAA 301
Qy 646 TCAGAAATAATATTTACTGAGAGATTAAAGGAGAGACAGTATTCTCGTCTCTCTGTT 705

```

```

Db 302 TCAGAAATATATTTTCGGGAAAGTTAAAGAGACTGTATTCTCGCACCTCGCTCTGTT 361
Qy 706 GGCAGTATTAAATCACTTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTGATCAAA 765
Db 362 GGAATATTAAAAATCAACTTCACCCCTCGAGTATTTCCCAACTGCACCTCCGTTGAATCAAAA 421
Qy 766 GTAGCAGAAGAGGAGAGTGGCTACACAAAACAAGCTGAGGCAGGACGACGAATGAATACT 825
Db 422 GTAGCAGAAGAGAGAGTGGCTACATAAACAGCTGAGGCAGGACGAGAAAGCAATGAATACT 481
Qy 826 GACATAGCTGCACTTTTCGGAATTTAAAAAGAGAGAGAAAGAACCCAGAAATGGTTGAAGAT 885
Db 482 GATATTCTCTGAATTTTTCGCACTTTAAAAAGAGAGAGAAAGAACCCAGAAATGGTTGAAGAG 541
Qy 886 AAAGGAAACAAATTTGTTGCAACGGAACATTTTGGCAGCTATCATGATATCAATAATTTA 945
Db 542 AAAGGAAACAAATTTGTTGCAACAGAGAGACTATTTGGCAGCTATTAAATGATACAAATTTA 601
Qy 946 GCCATAAGACTTAAATAAAGATGCCACTATTGTTATTTGAACCGGGCTGCTTGGCACCTA 1005
Db 602 GCCATAAGACTTAAATAAAGATTTCTTATTTATTTGAATCGGGCTGCTTGGCACCTA 661
Qy 1006 AAACATAAAAACTTTACACAAGGCTATTGAAAGATTTCTTAAAGCACTGGAATTTATGATG 1065
Db 662 AAACATAAAAACTTTACACAAGGCCATTGAAAGATTTCTTAAAGGCTCTGGAATTTACTGACA 721
Qy 1066 CCACCTGTTACAGACATGCTTAATGCAAGATGAGGACACATGTACCAGCTGGACAGCA 1125
Db 722 CCACATGTTGACAGCAATGCTTAATGCAAGATGAGGACACATGTACCAGCTGGACAGCA 781
Qy 1126 TTCTGTCAACT 1136
Db 782 TTCTGTCAACT 792

RESULT 5
LOCUS             BG771796          792 bp  mRNA  linear  EST 15-MAY-2001
DEFINITION        602720472P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837505 5',
                  mRNA sequence.
ACCESSION         BG771796
VERSION           BG771796.1 GI:14082449
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE          1 (bases 1 to 792)
AUTHORS            NIH-MGC http://mgc.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Straubeberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
                  cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
                  Toshiyuki and Piero Carninci (RIKEN)
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM10769 Row: p column: 18
                  High quality sequence stop: 671.
FEATURES          Location/Qualifiers
                    1..792
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:4837505"
                        /lab_host="DH10B"
                        /clone_lib="NIH_MGC_97"
                        /note="Organ: testis; Vector: pBluescriptR (modified

```


pBluescript KS+; Site 1: BamHI; Site 2: Sall-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN

Query Match 44.9%; Score 567.4; DB 4; Length 792;
Best Local Similarity 91.5%; Pred. No. 5, 5e-113;
Matches 671; Conservative 0; Mismatches 46; Indels 16; Gaps 6;
QY 1 ATGCTCTTTCAGGTTAGCGATTACAGTGGCAGACAGCAAGACTGCGGTCTTTCTGTCT 60
DB 37 ATGCTCTTTCAGGTTAGCGATTACAGTGGCAGACAGCAAGACTGCGGTCTTTCTGTCT 96
QY 61 CTGCCCCCTCAAGGGCGTGTGCTCAGACAGCAGGAGCTGTTCTGACGGAAAACTATCTG 120
DB 97 CTGCCCCCTCAAGGGCGTGTGCTCAGACAGCAGGAGCTGTTCTGACGGAAAACTATCTG 156
QY 121 AAGGTCAACTTTCCTCCATTTTATTGAGGCAATTTCTTTATGCTCCCATAGACCATGAG 180
DB 157 AAGGTCAACTTTCCTCCATTTTATTGAGGCAATTTCTTTATGCTCCCATAGACCATGAG 216
QY 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTTTCTTACCTTGTATATAAAAAAGAGCG 240
DB 217 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTTTCTTACCTTGTATATAAAAAAGAGCG 276
QY 241 GCCTGTGGGAGACCTTCTTGTGACGGGTGTGCAAGAGATGATGCAAGAAATTAGA 300
DB 277 GCCTGTGGGAGACCTTCTTGTGACGGGTGTGCAAGAGATGATGCAAGAAATTAGA 336
QY 301 GAAATATCTATTTCACAGCAAGAGAGAGCAAGAGAGCTACAGAGCAAGCAAGCTCA 360
DB 337 GAAATATCTATTTCACAGCAAGAGAGAGCAAGAGAGCTACAGAGCAAGCAAGCTCA 396
QY 361 GCAAGCGGAGAGATCAAAAAATPACGACTAAGTGTGATGATGATGATGATGATGATGATGAT 420
DB 397 GCAAGCGGAGAGATCAAAAAATPACGACTAAGTGTGATGATGATGATGATGATGATGATGAT 456
QY 421 AGGAAAAA-ATAGAAGATATGAAGAAAAATGAACGGATTAAGCCATTAAGCAATTGGA 479
DB 457 AGGAAAAAATAGAGATATGAAGAAAAATGAACGGATTAAGCCATTAAGCAATTGGA 516
QY 480 AGCCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAGAAAGAA--TTGAGAGAGAG 537
DB 517 AGCCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAGAAAGAA--TTGAGAGAGAG 576
QY 538 AAATATGTCAAAAAGA-----AAAGCAAAATTAAGAGAGGAGAAAAAATAAATAAT 591
DB 577 AAATATGTCAAAAAGAAAGCAAAATTAAGAGAGGAGAAAGACATACCAATACAAATAT 636
QY 592 AAGAGCTTACTAGAAAATTTGGCATCTAGAAATCTTG---CTCCAAAAGGGAGAAATTC 648
DB 637 AAGAGCTTACTAGAAAATTTGGCATCTAGAAATCTTGCTCCAGAGGGGAGAAATTC 696
QY 649 G--AAAATATATTACTGAGAAATTAAGAGAGACAGTATTC--TGCTCCTCGCTCTGT 704
DB 697 GCCCAATATATTACTGAGAGATTAAGAGAGAAACAGTATTCCTCGCTCGCTCTGT 756
QY 705 TGGCAGTATTA 717
DB 757 TGGCAGACTTAA 769

RESULT 6

BE564350
LOCUS 601343161F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3685335 5',
DEFINITION mRNA sequence.

ACCESSION

BE564350

BE564350.1 GI:9808070

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..788
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3685335"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 53"
/note="Organ: bladder; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.55
kb (range 0.9-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

ORIGIN

Query Match 43.8%; Score 553; DB 2; Length 788;
Best Local Similarity 94.6%; Pred. No. 7, 6e-110;
Matches 594; Conservative 0; Mismatches 31; Indels 3; Gaps 2;
QY 1 ATGCTCTTTCAGGTTAGCGATTACAGTGGCAGACAGCAAGACTGCGGTCTTTCTGTCT 60
DB 23 ATGCTCTTTCAGGTTAGCGATTACAGTGGCAGACAGCAAGACTGCGGTCTTTCTGTCT 82
QY 61 CTGCCCCCTCAAGGGCGTGTGCTCAGACAGCAGGAGCTTCTGCACGAAAACTATCTG 120
DB 83 CTGCCCCCTCAAGGGCGTGTGCTCAGACAGCAGGAGCTTCTGCACGAAAACTATCTG 142
QY 121 AAGGTCAACTTTCCTCCATTTTATTGAGGCAATTTCTTTATGCTCCCATAGACCATGAG 180
DB 143 AAGGTCAACTTTCCTCCATTTTATTGAGGCAATTTCTTTATGCTCCCATAGACCATGAG 202
QY 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTACCTTGTATATAAAAAAGAGCG 240
DB 203 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTACCTTGTATATAAAAAAGAGCG 262
QY 241 GCCATGTGGGAGACCTTTCGTGACGGGTGTGCAAGAGAGATGATGCAAGAAATTAGA 300
DB 263 GCCATGTGGGAGACCTTTCGTGACGGGTGTGCAAGAGAGATGATGCAAGAAATTAGA 322
QY 301 GAAAAATCTATTTCACAGCAAGAGAGAGCAAAAAAGAGCTACAGAGCAAGAAAGCTGCA 360
DB 323 GAAAAATCTATTTCACAGCAAGAGAGAGCAAAAAAGAGCTACAGAGCAAGAAAGCTGCA 382
QY 361 GCAAGCGGGAAGATCAAAAAATACGCATAAGTGTGATGATGATGATGATGATGATGATGAT 420

```

Db      383  GCAAGCGGAGATCAAAATACGCACTAAGTCTCATGATGAAGATTGACGCGAAGAG 442
Qy      421  AGG--AAAAAATAGAGATATGAAGAAATGAACCGGATATAAAGCCACTAAAGCATTTGG 478
Db      443  AGGCAAAACAAATAGAGATATGAAGAAATGAACCGGATATAAAGCCACTAAAGCATTTGG 502
Qy      479  AAGCCTGGAGATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAAAGAGA 538
Db      503  AAGCCTGGAAGAATATCAAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAAAGAGA 562
Qy      539  AATTATGTCAAAAAAGAAAGCAAAATTAAGAAAGCAAAAAAATTAATAATATAGAGTC 598
Db      563  ACTTTGTC-GACAGAAAGCACTTCACGAGGCGGACACACCAACATTAATATAGAGTC 621
Qy      599  TTACTAGAAATTTGGCATCTAGAAATCT 626
Db      622  TTAAGAAATGTGGCTTCTAGAAATCT 649

```

```

RESULT 7
BP460416 683 bp mRNA linear EST 31-DEC-2003
LOCUS BP460416 full-length enriched swine cDNA library, adult ovary Sus
DEFINITION BP460416 cDNA clone OVRM10147G01 5', mRNA sequence.
ACCESSION BP460416
VERSION BP460416.1 GI:40476478
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

```

```

REFERENCE 1
AUTHORS Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H., Okumura,N., Hamsima,N. and Awata,T.
TITLE PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries
JOURNAL Nucleic Acids Res. 32 (1), D484-D488 (2004)
COMMENT Contact: Hirohide Uenishi
Animal Genome Laboratory, Genome Research Department
National Institute of Agrobiological Sciences
2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
Tel.: +81-29-838-8627
Fax: +81-29-838-8627
Email: huenishi@affrc.go.jp
EST project with full-length enriched cDNA libraries carried out in
Animal Genome Research Program (Japan) by National Institute of
Agrobiological Sciences and STAFF-Institute
Single pass sequencing of clones derived from oligo-capped cDNA
library
Vector sequences were eliminated by RepeatMasker version 2002/07/13
and crossmatch version 0.990319
Low quality bases were trimmed based on the quality values.

```

```

FEATURES             source
Location/Qualifiers
1..683
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="OVRM10147G01"
/tissue_type="ovary"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult ovary"

```

```

ORIGIN
Query Match 41.1%; Score 519; DB 5; Length 683;
Best Local Similarity 90.2%; Pred. No. 1.9e-102;
Matches 555; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy      539  AATTATGTCAAAAAAGAAAGCAAAATTAAGAAAGGAGAAAAAATAATAATATAGAGTC 598
Db      1    AATTACATCAACGAAGCAAAATTTGAAGAGAGAGAAAAAATTAAGCGCTAAAGGCC 60

```

```

Qy      599  TTACTAGAAATTTGGCATCTAGAAATCTTCTGCCAAAGGGAGAGAAATTCAGAAATATAT 658
Db      61    TTACTAGAAATTTGAGCATCTAGATATCTTGACACAAATGGAGAAATTCAGAAATATAT 120
Qy      659  TTACTGAGAGTTAAAGGAAGACAGTATCTCTGCTCTCTGCTCTGCTGCGAGTATTTAAAA 718
Db      121  TTTTGAAGAGTTAAAGGAAGACAGTATCTCTGCTCTCTGCTCTGCTGCGAGTATTTAAAA 180
Qy      719  TCACTTTTACCCCTCGAGTATTTCCACACAGCTCTTCTGCTGAATCACAAGTAGCAGCAAGG 778
Db      181  TCAACTTTTACCCCTCGAGTATTTCCCAACCGCCCTCCCGGGAATCACAAGTAGCGGAAGAAG 240
Qy      779  AGAGTGGCTACACAAACAAAGCTGAGGCAGCAAGAGCAATGAATACATCAGATAGCTGAAC 838
Db      241  AGAGTGGCTACACAAACAAAGCAGAGGCAAGAGGGCAATGAATCTGAAATTCCTGAGT 300
Qy      839  TTTGCGATTTAAAGGAAGAAAGAAACCCAGAAATGGTTGAAGGATAAAGGAACAAAT 898
Db      301  TTAGTGATTTAAAGGAAGAAAGAAACCCAGAAATGGTTGAAGGACAAAGAACCAAGT 360
Qy      899  TGTTTGCAACGGAAACTATTTGCGAGCTATCAATGCATATATTTAGCCATAGACTAA 958
Db      361  TGTTTGCAACGAGAAACTATTTGCGAGCTATTAATGCATACAACTTAGCCATAGACTAA 420
Qy      959  ATAATAAGATGCCACTATTTGTTGAACCGGGCTGCTTGCCACTAAAACTAAAAAACC 1018
Db      421  ATAATAAGATTCACACTGTTGTTGATCGGGCTGCTTGCCACTAAAACTAAAAAACC 480
Qy      1019  TACACAAGGCTATTGAAGATTTCTTAAGGCACCTGGAATTTATGATGCCACCTGTTACAG 1078
Db      481  TACACAAGGCTATTGAAGATTTCTTAAGGCATTTAGAATTTATTAACACCACTGTTGAG 540
Qy      1079  ACATGCTTAATGCAGAAATGAAGGCACATGTACGACCTGGAACAGCATTTCTGCAACTAG 1138
Db      541  ACAATGCTTAATGCAGAAATGAAGGCACATGTACGACCTGGAACAGCATTTCTGCAACTAG 600
Qy      1139  AATTGTATGTAGAAG 1153
Db      601  AATTGTATGTAGAAG 615

```

```

RESULT 8
LOCUS AK005832
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700010i24 product:EKX1 homolog [Homo sapiens], full insert sequence.
ACCESSION AK005832
VERSION AK005832.1 GI:12838612
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

```

```

REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

```


LOCUS BQ217312 1131 bp mRNA linear EST 02-MAY-2002
 DEFINITION AGENCOURT_7558431 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6047052
 5', mRNA sequence.
 ACCESSION BQ217312
 VERSION BQ217312.1 GI:20398712
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCFT/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM13293 row: n column: 13
 High quality sequence stop: 398.
 Location/Qualifiers
 1. .1131
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6047052"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 ORIGIN
 Query Match 40.4%; Score 510.4; DB 5; Length 1131;
 Best Local Similarity 95.9%; Pred. No. 1.5e-100;
 Matches 535; Conservative 0; Mismatches 21; Indels 2; Gaps 1;
 QY 42 GACTGCGGCTTTCTGCTCTGCTGCCCTCAAAGGCGGTGCGTCAGACACGGACGGTGT 101
 DB 1 GACTGCGGCTTTCTGCTCTGCTGCCCTCAAAGGCGGTGCGTCAGACACGGACGGTGT 60
 QY 102 CTGCACGGAAACTATCTGAGGGTCACTTTCCTCCATTTTATTGAGGCATTTCTTTA 161
 DB 61 CTGCACGGAAACTATCTGAGGGTCACTTTCCTCCATTTTATTGAGGCATTTCTTTA 120
 QY 162 TGCTCCCATAGACATGAGCAGCAAGCAAGCAAGATGGGAATGACACCATTTGCTTCAC 221
 DB 121 TGCTCCCATAGCAGTGAAGCAGCAAGCAAGATGGGAATGACACCATTTGCTTCAC 180
 QY 222 CTTGTATATAAAGAACGGCCATGTGGAGACCCCTTTCTGTGACGGGTGTGCAAGA 281
 DB 181 CTTGTATATAAAGAACGGCCATGTGGAGACCCCTTTCTGTGACGGGTGTGCAAGA 240
 QY 282 GATGATGCAAGATATAGAGAAATCTATTTTACAAGCACAAGAGAGCAAAAGAGC 341
 DB 241 GATGATGCAAGATATAGAGAAATCTATTTTACAAGCACAAGAGAGCAAAAGAGC 300
 QY 342 TACAGACCAAAAGCTGCAGCAAAAGCGGAAGATCAAAATAACGACCTAAGTGTGATCAT 401
 DB 301 TACAGACCAAAAGCTGCAGCAAAAGCGGAAGATCAAAATAACGACCTAAGTGTGATCAT 360
 QY 402 GAAGATTGAAGAAAGAGAGAGAAATAATGAGATATGAAGAAATGAACGGATAAA 461
 DB 361 GAAGATTGAAGAAAGAGAGAGAAATAATGAGATATGAAGAAATGAACGGATAAA 420
 QY 462 AGCCACTAAGCATTTGGAAGCTGGAAGAAATATCAAGAAAGCTGAGGAGC--AAAAA 519

Db 421 AGCCACTAAGCATTTGGAAGCTGGAAGAAATATCAAGAAAGCTGAGGAGCAAAAAA 480
 QY 520 AAAATTCAGAGAGAGAGAAATTTATGTCAAAAGAGAAAGCAAAATTAAGAGAGGAGAAA 579
 Db 481 AAAATTCAGAGAGAGAGAAATTTATGTCAAAAGAGAAAGCAAAATTAAGAGAAATGAAGGA 540
 QY 580 AAAATTAATAATTAAGACT 597
 Db 541 AAAAAAATAATAATAGT 558
 RESULT 10
 LOCUS BY705981
 DEFINITION BY705981 RIKEN full-length enriched, adult male testis Mus musculus
 cDNA clone 1700010124 5', mRNA sequence.
 ACCESSION BY705981
 VERSION BY705981.1 GI:27117128
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 917)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gustincich, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedziera, R.M., King, B.L., Konagaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takeru, Y., Taylor, M.S., Teasdale, R., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, D., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
 Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
 Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
 Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
 Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission


```

QY      866 ACCC 869
DB      514 AGCC 517

RESULT 13
LOCUS   CK603982
DEFINITION AGENCOURT_17890681 NIH_MGC_238 Rattus norvegicus cDNA clone
IMAGE: 7134374 5', mRNA sequence.
ACCESSION CK603982
VERSION   CK603982.1 GI:41117319
KEYWORDS  EST.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 796)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Howard Jacobs
          cDNA Library Preparation: Express Genomics
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM15039 row: o column: 12
          High quality sequence stop: 622.
          Location/Qualifiers
            1..796
              /organism="Rattus norvegicus"
              /mol_type="mRNA"
              /db_xref="taxon:10116"
              /clone="IMAGE:7134374"
              /tissue_type="testis, pooled"
              /lab_host="DH10B Tona"
              /clone_lib="NIH MGC 238"
              /note="Organ: testis; Vector: pExpress-1; Site 1: EcoRV;
              Site 2: NotI; RNA obtained from testis tissue of 8 wk old
              animal. Tissues were snap-frozen and kept at -80C before
              RNA extraction and purification (tri-reagent method). cDNA
              was primed using oligo-dT primer:
              5'-pGACTAGTCTAGATCGGCGGCGGCC(T)25-3' and cloned into
              the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
              resulted in an average insert size of 1.9 kb. This primary
              library is normalized (non-normalized primary library is
              NIH MGC 237) and was constructed by Express Genomics
              (Frederick, MD)"

FEATURES             source
    source
      Query Match      35.6%; Score 449; DB 7; Length 796;
      Best Local Similarity 76.9%; Pred. No. 3.4e-87;
      Matches 574; Conservative 0; Mismatches 166; Indels 6; Gaps 2;

QY      1 ATGCGCTCTTCAGGTAGCATTAACAGTGGCAGACGAGACGAGCTGGCTCTTTCTGTCT 60
DB      |||||
DB      4 ATGCGGGTGGAGTAGCGAGCTAGCTGGGAGCAGACACCGCGCGCACTCTTCTGTGCG 63
QY      61 CTGCGCCCTCAAAAGGCGTGTGCTCAAGACACGAGCGTGTCTGCAACGAAACATATCTG 120
DB      |||||
DB      64 CTGCGCTCTACGGGGGCTGCTGCGCGCATGTGATGTTCTGTGGGGAAGTTACCTG 123
QY      121 AAGGTCAACTTTCCTCCATTTTATTTAGGCAATTCCTTATGCTCCCATAGACCATGAG 180
DB      |||||
DB      124 AAGGTTAACTTTCCTCCATTTTATTTAGGAGTGTTCCTATGCTCCCATAGATGAGG 183

```

```

QY      181 AGCAGCAAAAGCAAGATTGGGAATGACACCAATTGTCTTCACTTGTATATAAAAGAGCG 240
DB      |||||
DB      184 AAGAGCAAAAGCCAGATTGGAAATGACACGATTCTCTTCACTTGTATATAAAAGAGGCCA 243
QY      241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATGCARAGAATTAGA 300
DB      |||||
DB      244 GTTCTGTGGGAGAGCCCTTTCTATGACCGCGTTGTATTAAGAGATGATGCAGAGATAAGA 303
QY      301 GAAATAATCTATTTTACAAGCAACAAGAGAGAGCAAAAGAGCTTACAGAAGCAAAAGCTGCA 360
DB      |||||
DB      304 GAAAATCTATCTTGTCAAGACAGAGAGAAAGCAAAAGAGGCTTACAGAGCGAAGCTGCT 363
QY      361 GCAAGAGCGGAGAGATCAAAAATACCCACTAAGTGTGATGATGAAGATTGAAGAAGAAGAG 420
DB      |||||
DB      364 GCCAAGCCGAGAAGATCAGAGATACGCCCTTAGCCGAGATGATGAAGATTGAAGAGGAAGAG 423
QY      421 AGGAAAAAATAGAGATATGAAGAAATGAACCGATGAACCGATGAACCGATGAACCGATGAAC 480
DB      |||||
DB      424 AGGAAAAAATAGAGATATGAAGAAATGAACCGATGAACCGATGAACCGATGAACCGATGAAC 483
QY      481 GCCTGGAAGAAATATCAAGAGAAAGCTGAGGAGCAAAAAAATTCAGAGAGAGAGAGAAA 540
DB      |||||
DB      484 GCGTGAAGAAATGCCAAAGAAAGCTGACGGAACAAAACGAGTCCAGAGGAGAGAGAA- 542
QY      541 TTATGTCAAAAAGAAAGCAAAATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGTCTT 600
DB      |||||
DB      543 --ACCGCTACAGGGGAAAGCAAGCT--GAAGAGAGGGGAGGCACTAAAACCTCAGAGTTTG 597
QY      601 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB      |||||
DB      598 CCCCAGAGGGCCCCGCCCACTCGCCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 657
QY      661 ACTGAGAGAGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB      |||||
DB      658 TCTGAGAGAGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
QY      721 AACTTTACCCCTCGAGATATCCCAAC 746
DB      |||||
DB      718 AGCTTTNACCCTCGAGTGTCCCAAC 743

RESULT 14
LOCUS   CB297042
DEFINITION CB297042.1 GI:28622472
ACCESSION CB297042
VERSION   CB297042.1
KEYWORDS  EST.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 468)
AUTHORS   Hellmann,I., Zollner,S., Enard,W., Ebersberger,I., Nickel,B. and
          Paabo,S.
TITLE     Selection on human genes as revealed by comparisons to chimpanzee
JOURNAL   Genome Res. (2003) In press
COMMENT   Contact: Paabo S
          Evolutionary Genetics
          Max-Planck-Institute for evolutionary Anthropology
          Deutscher Platz 6, 04103 Leipzig, Germany
          Tel: +49-(0)-341-3550 500
          Fax: +49-(0)-341-3550 555
          Email: paabo@eva.mpg.de
          Seq primer: M13 reverse.
          Location/Qualifiers
            1..468
              /organism="Pan troglodytes"
              /mol_type="mRNA"

```


Search completed: July 8, 2005, 04:46:16
Job time : 4610 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:40:53 ; Search time 245 Seconds
(without alignments)
8435.173 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaaactaaattcttaa 1263

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	
1	432.6	34.3	489	US-09-621-976-1234	Sequence 1234, Ap	
2	178.8	14.2	313	US-09-513-999C-30874	Sequence 30874, A	
3	162.8	12.9	164	US-09-513-999C-34573	Sequence 34573, A	
C	4	115.8	9.2	7218	US-08-232-463-14	Sequence 14, Appl
5	90.8	7.2	612	US-09-902-540-1357	Sequence 1357, Ap	
6	77.4	6.1	1039	US-09-902-540-1280	Sequence 1280, Ap	
7	77.2	6.1	614	US-09-902-540-1318	Sequence 1318, Ap	
8	73.4	5.8	258775	US-09-949-016-16435	Sequence 16435, A	
9	72.6	5.7	705	US-09-270-767-5061	Sequence 5061, Ap	
10	72.6	5.7	705	US-09-270-767-20343	Sequence 20343, A	
11	72.4	5.7	240	US-08-628-417-6	Sequence 6, Appl	
C	12	72	5.7	396	US-09-640-173-53	Sequence 53, Appl
13	72	5.7	396	US-09-713-550-53	Sequence 53, Appl	
C	14	72	5.7	396	US-09-825-294-53	Sequence 53, Appl
C	15	72	5.7	396	US-09-970-966-53	Sequence 53, Appl
16	71.2	5.6	1696	US-09-835-811-1	Sequence 1, Appl	
17	71.2	5.6	1827	US-09-270-767-1308	Sequence 1308, Ap	
18	71.2	5.6	1827	US-09-270-767-16590	Sequence 16590, A	
C	19	70.2	5.6	851	US-09-270-767-8286	Sequence 8286, Ap
C	20	70.2	5.6	851	US-09-270-767-23568	Sequence 23568, A
21	69.8	5.5	161124	US-09-949-016-11760	Sequence 11760, A	
22	69.4	5.5	2447	US-09-014-969-14	Sequence 14, Appl	
C	23	67.4	5.3	25431	US-09-949-016-13234	Sequence 13234, A
24	67.2	5.3	601	US-09-949-016-184917	Sequence 184917, A	
25	67.2	5.3	601	US-09-949-016-185072	Sequence 185072, A	
26	66.6	5.3	194937	US-09-949-016-17032	Sequence 17032, A	
27	66.6	5.3	194937	US-09-949-016-17033	Sequence 17033, A	

28	66.2	5.2	696	3	US-09-461-697-193	Sequence 193, App	
29	66.2	5.2	699	3	US-09-461-697-191	Sequence 191, App	
30	66.2	5.2	717	3	US-09-461-697-189	Sequence 189, App	
31	66.2	5.2	774	3	US-09-461-697-187	Sequence 187, App	
32	66.2	5.2	819	3	US-09-461-697-185	Sequence 185, App	
33	66.2	5.2	1669	3	US-09-461-697-184	Sequence 184, App	
34	66.2	5.2	19124	2	US-08-487-826B-13	Sequence 13, Appl	
35	65.4	5.2	36731	4	US-09-949-016-13770	Sequence 13770, A	
36	64.6	5.1	2223	1	US-08-257-073-4	Sequence 4, Appl	
37	63.6	5.0	980	3	US-09-171-209-8	Sequence 8, Appl	
C	38	63.6	5.0	55841	4	US-09-949-016-16602	Sequence 16602, A
39	63.6	5.0	74790	4	US-09-949-016-15321	Sequence 15321, A	
40	63.4	5.0	5433	3	US-08-929-329-1	Sequence 1, Appl	
C	41	63.4	5.0	12703	4	US-09-949-016-16685	Sequence 16685, A
42	62.6	5.0	237510	4	US-09-949-016-14273	Sequence 14273, A	
C	43	62.4	4.9	54033	4	US-09-949-016-12091	Sequence 12091, A
C	44	62.4	4.9	54033	4	US-09-949-016-14325	Sequence 14325, A
C	45	62	4.9	1298	3	US-08-948-705-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-621-976-1234
; Sequence 1234, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1234
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..488
US-09-621-976-1234

Query Match	34.3%	Score 432.6;	DB 4;	Length 489;
Best Local Similarity	99.8%	Pred. No. 4.6e-97;		
Matches	432;	Conservative 1;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGAAAGACTGCGGTCTTTCTGTCT	60	
Db	57	ATGCTCTTTCAGGTTAGCGATTACAKTGGCAGCAGACGAAAGACTGCGGTCTTTCTGTCT	116	
Qy	61	CTGCCCTCAAAAGCGGTGGCTCAGACACGACGCTGTTCTGCAGGAAACTATCTG	120	
Db	117	CTGCCCTCAAAAGCGGTGGCTCAGACACGACGCTGTTCTGCAGGAAACTATCTG	176	
Qy	121	AAGTCAACTTTCCTCCATTTTATTTAGGCATTTCTTTATGTCCCATAGACGATCAG	180	
Db	177	AAGTCAACTTTCCTCCATTTTATTTAGGCATTTCTTTATGTCCCATAGACGATCAG	236	
Qy	181	AGCAGCAAGCAAGATTGGGAATGACACCACTTCTTTCACCTTTGTATATAAAGAACGCG	240	
Db	237	AGCAGCAAGCAAGATTGGGAATGACACCACTTCTTTCACCTTTGTATATAAAGAACGCG	296	
Qy	241	GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATAGA	300	
Db	297	GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATAGA	356	
Qy	301	GAAAAATCTATTTTCAAGCACAAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA	360	
Db	357	GAAAAATCTATTTTCAAGCACAAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA	416	

Qy	361	GCAAAGCGGAAGATCAAAAAATACGCCTAAGTGTCTATGATGAAGATTGAAGAAGAAGAG	420
Db	417	GCAAAGCGGAAGATCAAAAAATACGCCTAAGTGTCTATGATGAAGATTGAAGAAGAAGAG	476
Qy	421	AGGAAAAAATAG	433
Db	477	AGGAAAAAATAG	489

```

RESULT 2
US-09-513-999C-30874
; Sequence 30874, Application US/09513999C
;
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

```

Query Match	14.2%	Score 178.8;	DB 4;	Length 313;
Best Local Similarity	94.8%	Pred. No. 1.5e-34;		
Matches 201; Conservative	5;	Mismatches 3;	Indels 3;	Gaps 2;

844	Qy	GATTTAAAGAAAGAAAAGAACCCAGAAATGGTTGAAGGATAAAGGAAACAAATGTGTT	903
1	Db	GATTTAAAGAAAGAAAAGAACCCAGAAATGGTTGAAGGATAAAGGAAACAAATGTGTT	60
904	Qy	GCAACGGAAAACTATTTCGCAGCTA--TCAATGCATAT-AAATTAGCCATAAGACTAAAT	960
61	Db	GCAACGGAAAACTATTTCGCARCWATCAVAAWGCTATAAAATTAGCCATAAGACTAAAT	120

Qy	961	ATAAAGTGGCACTATTGTATTTGAACCGGGCTCTTGCCACCTTAAAACTTAAAAACTTAA	102
Db	121	ATAAAGATGCCATTATTGTATTTGAACCGGGCTCTTGCCACCTTAAAACTTAAAAACTTAA	180
Qy	1021	CACAAGCTATTGAAGATTTCTTCAAGGCAC	1052
Db	181	CACAAGCTATTGAAGATTTCTTCAAGGAATT	212

RESULT 3
US-09-513-999C-34573
; Sequence 34573, Application US/09513999C
; Patent NO. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Query Match	12.9%	Score 162.8;	DB 4;	Length 164;
Best Local Similarity	98.8%	Pred. No. 9.9e-31;		
Matches 162:	Conservative	2: Mismatches	0: Indels	0: Gaps

[illegible]

```

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA

```



```
; MOLECULE TYPE: oligodeoxynucleotide
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-628-417-6

Query Match          5.7%; Score 72.4; DB 1; Length 240;
Best Local Similarity 57.5%; Pred. No. 2.2e-08;
Matches 130; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 370 GAAGATCAAAATACGCACTAAGTGTGTCATGATGAAGATTGAAGAAGAGAGAGAAAAA 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11 GACAATAAACTTTAGAAATAATTTTACTAAAAAAGAGAGAGAGAGAGAGAGAGAGAG 70

Qy 430 ATAGAAGTATGAAGAAGAAATGACGCTAAAGCCCTAAAGCATTTGGAAGCTGGAAA 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 130

Qy 490 GAATATCAAGAAAGCTGAGGAGCAAAAAAAAAAAAAAAAAATTCAGAGAGAGAGAAATTTATGTCAA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 190

Qy 550 AAGAAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 595
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 236

RESULT 12
US-09-640-173-53/c
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; FILE REFERENCE: METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640.173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

Query Match          5.7%; Score 72; DB 4; Length 396;
Best Local Similarity 54.9%; Pred. No. 3.4e-08;
Matches 135; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy 363 AAAGCGGAGATCAAAATAACGCACTAAGTGTGTCATGATGAAGATTGAAGAAGAGAGAG 422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 ANAAGGNAANNAANNAANNTAAANNAANNAANNAANNAANNAANNAANNAANNA 188

Qy 423 GAAAAAATGACGATATGAAGAAGAAATGAACGGATTAAGCCCTAAAGCATTTGGAAGC 482
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

Qy 483 CTGGAAGATATCAAGAGAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAGAGAAATT 542
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68

Qy 543 ATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

Qy 603 TAGAAA 608
    |||||
Db 7 AAAAAA 2

RESULT 13
US-09-713-550-53/c
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713.550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match          5.7%; Score 72; DB 4; Length 396;
Best Local Similarity 54.9%; Pred. No. 3.4e-08;
Matches 135; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy 363 AAAGCGGAGATCAAAATAACGCACTAAGTGTGTCATGATGAAGATTGAAGAAGAGAGAG 422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 ANAAGGNAANNAANNAANNTAAANNAANNAANNAANNAANNAANNAANNAANNA 188

Qy 423 GAAAAAATGACGATATGAAGAAGAAATGAACGGATTAAGCCCTAAAGCATTTGGAAGC 482
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

Qy 483 CTGGAAGATATCAAGAGAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAGAGAAATT 542
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68

Qy 543 ATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

Qy 603 TAGAAA 608
    |||||
Db 7 AAAAAA 2

RESULT 14
US-09-825-294-53/c
; Sequence 53, Application US/09825294
; Patent No. 6710170
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C5
; CURRENT APPLICATION NUMBER: US/09/825.294
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)....(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53

Query Match          5.7%; Score 72; DB 4; Length 396;
Best Local Similarity 54.9%; Pred. No. 3.4e-08;
Matches 135; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 363 AAAGCGGAAGATCAAAATACGCCTAAGTCTCATGTGAAGATTGAAGAGAGAGAG 422
Db 247 ANAAGGNAAAANAAAANNTAAANNTAAANNTAAANNTAAANNTAAANNTAAANNT 188
QY 423 GAAAGGATAGAGATATCAAGAGAAATGAACGGATAAAGCCACTTAAGCATTGGAAGC 482
Db 187 AAAAAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 128
QY 483 CTGGAAGAGATATCAAGAGAAAGCTGAGGAGCAAAAGAAATTCAGAGAGAGAGAG 542
Db 127 AAAAAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 68
QY 543 ATGTCAGAAAGGAAAGCAAAATTAAGAGAGGAAAGAAATTAAGAGAGGAGAGAG 602
Db 67 AAAAAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 8
QY 603 TAGAAA 608
Db 7 AAAAAA 2

Search completed: July 8, 2005, 04:50:19
Job time : 248 secs
```

```
RESULT 15
US-09-970-966-53/c
; Sequence 53, Application US/09970966
; Patent No. 6720146
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesch, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 224, 225, 228, 235, 240, 246, 257, 266, 274, 279, 281, 282,
; LOCATION: 283, 285, 287, 288, 290, 291, 292, 293, 294, 295, 296, 297,
; LOCATION: 300, 301, 303, 307, 311, 313, 314, 317, 318, 319, 320, 321,
; LOCATION: 323, 324, 328, 329, 330, 336, 337, 338, 339, 340, 341
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 355,
; LOCATION: 357, 358, 359, 362, 363, 364, 365, 366, 367, 373, 380, 381,
; LOCATION: 382, 385, 387, 388, 389, 390, 392
; OTHER INFORMATION: n = A,T,C or G
US-09-970-966-53

Query Match          5.7%; Score 72; DB 4; Length 396;
Best Local Similarity 54.9%; Pred. No. 3.4e-08;
Matches 135; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 363 AAAGCGGAAGATCAAAATACGCCTAAGTCTCATGTGAAGATTGAAGAGAGAGAG 422
Db 247 ANAAGGNAAAANAAAANNTAAANNTAAANNTAAANNTAAANNTAAANNTAAANNT 188
QY 423 GAAAGGATAGAGATATCAAGAGAAATGAACGGATAAAGCCACTTAAGCATTGGAAGC 482
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 03:29:26 ; Search time 1941 Seconds
(without alignments)
4085.271 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgctcttcaggtagcgca.....gaacagaactaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	1263	17	US-10-364-505-1
2	1263	100.0	1263	19	US-10-681-199-1
3	1263	100.0	1993	17	US-10-364-505-2
4	1263	100.0	1993	19	US-10-681-199-2
5	1263.4	99.2	1263	19	US-10-681-199-19
6	1251.8	99.1	1263	19	US-10-681-199-13
7	1248.6	98.9	1263	19	US-10-681-199-15

8	1243.8	98.5	1263	19	US-10-681-199-17	Sequence 17, Appl
9	1047.4	82.9	1559	17	US-10-108-260A-575	Sequence 575, Appl
10	840.8	66.6	1697	17	US-10-364-505-4	Sequence 4, Appli
11	840.8	66.6	1697	19	US-10-681-199-4	Sequence 4, Appli
12	401.6	31.8	464	17	US-10-242-535A-30001	Sequence 30001, A
13	401.6	31.8	464	18	US-10-085-783A-30001	Sequence 30001, A
14	369.8	29.3	488	10	US-09-918-995-22675	Sequence 22675, A
15	336	26.6	458	10	US-09-918-995-10303	Sequence 10303, A
16	325	25.7	325	18	US-10-240-425-872	Sequence 872, App
17	264.2	20.9	715	21	US-10-956-157-3114	Sequence 3114, Ap
18	264.2	20.9	715	21	US-10-956-157-8349	Sequence 8349, Ap
19	239.4	19.0	50000	17	US-10-364-505-7	Sequence 7, Appli
20	239.4	19.0	50000	19	US-10-681-199-7	Sequence 7, Appli
21	219.2	17.4	505	9	US-09-783-590-8049	Sequence 8049, Ap
22	156.6	12.4	50000	17	US-10-364-505-8	Sequence 8, Appli
23	156.6	12.4	50000	19	US-10-681-199-8	Sequence 77858, A
24	99	7.8	1204	19	US-10-437-963-77858	Sequence 1931, Ap
25	97.6	7.7	14006	15	US-10-311-455-1931	Sequence 1931, Ap
26	97.6	7.7	3673778	16	US-10-312-841-2	Sequence 2, Appli
27	96.8	7.7	16033	15	US-10-311-455-1377	Sequence 1377, Ap
28	95.4	7.6	803	19	US-10-437-963-72176	Sequence 72176, A
29	94.8	7.5	880	20	US-10-425-115-53133	Sequence 53133, A
30	94.2	7.5	1243	20	US-10-425-115-172717	Sequence 172717, A
31	94	7.4	921	20	US-10-425-115-38710	Sequence 38710, A
32	93.4	7.4	447	19	US-10-021-323-9573	Sequence 9573, Ap
33	93.4	7.4	986	20	US-10-425-115-174478	Sequence 174478, A
34	93.2	7.4	529	9	US-09-983-965-2109	Sequence 2109, App
35	93.2	7.4	2520	17	US-10-389-566-205	Sequence 205, App
36	93	7.4	1062	20	US-10-425-115-120013	Sequence 120013, A
37	92.8	7.3	627	19	US-10-021-323-9336	Sequence 9336, Ap
38	92.8	7.3	6668	15	US-10-311-455-1670	Sequence 1670, Ap
39	92.8	7.3	17934	15	US-10-311-455-1692	Sequence 1692, Ap
40	92.4	7.3	602	20	US-10-425-115-182618	Sequence 182618, A
41	92.2	7.3	556	20	US-10-363-345A-23587	Sequence 23587, A
42	92.2	7.3	556	21	US-10-363-345A-23588	Sequence 23588, A
43	92.2	7.3	556	21	US-10-363-483A-23587	Sequence 23587, A
44	92.2	7.3	556	21	US-10-363-483A-23588	Sequence 23588, A
45	91.4	7.2	1214	18	US-10-424-599-102083	Sequence 102083, A

ALIGNMENTS

RESULT 1
US-10-364-505-1
; Sequence 1, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
; APPLICANT: Kere, Juha
; APPLICANT: Taipale, Mikko
; APPLICANT: No. US20030219787A1ola-Hemmi, Jaana
; APPLICANT: Kaminen, Nina
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0199P
; CURRENT APPLICATION NUMBER: US/10/364.505
; CURRENT FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequence for human DYX1 (cDNA)

US-10-364-505-1

Query Match 100.0%; Score 1263; DB 17; Length 1263;
Best Local Similarity 100.0%; Pred. No. 5e-253;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTCTTCAGGTAGCGATTACAGTGGCAGCAGACGAGACTGGGCTTTCTTCTCT 60
|||||
Db 1 ATGCTCTTCAGGTAGCGATTACAGTGGCAGCAGACGAGACTGGGCTTTCTTCTCT 60


```
QY 541 TTATGTCACAAAGCAAGCAGATTTAAAGAGGAGCAAGAAAGAAATATAATATAGAGTCTT 600
Db 541 TTATGTCACAAAGCAAGCAGATTTAAAGAGGAGCAAGAAAGAAATATAATATAGAGTCTT 600
QY 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAGAAATTCAGAAATATATTT 660
Db 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAGAAATTCAGAAATATATTT 660
QY 661 ACTGAGAGTTAAAGAGACAGATATCTGCTCCTCGCTCTGTTGGCAGATATTAATATC 720
Db 661 ACTGAGAGTTAAAGAGACAGATATCTGCTCCTCGCTCTGTTGGCAGATATTAATATC 720
QY 721 AACTTTACCCCTCGAGTATTTCCACAGCTCTTCTGTAATCAAGTACAGAGGAG 780
Db 721 AACTTTACCCCTCGAGTATTTCCACAGCTCTTCTGTAATCAAGTACAGAGGAG 780
QY 781 GAGTGGCTACACAAACAGCTGAGGACGAAGAGCAATGAATPACTGACATAGCTGAAT 840
Db 781 GAGTGGCTACACAAACAGCTGAGGACGAAGAGCAATGAATPACTGACATAGCTGAAT 840
QY 841 TCGGATTTAAAGAGAGAAAGAACCCAGAAATGTTGAAGGATAAAGGAAACAAATG 900
Db 841 TCGGATTTAAAGAGAGAAAGAACCCAGAAATGTTGAAGGATAAAGGAAACAAATG 900
QY 901 TTTGCAACGGAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAGACTAAT 960
Db 901 TTTGCAACGGAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAGACTAAT 960
QY 961 AATAAGATGCCACTATTTGTAATTTGAACCGGGCTGCTGGCCACTTAAACTTAAACCTTA 1020
Db 961 AATAAGATGCCACTATTTGTAATTTGAACCGGGCTGCTGGCCACTTAAACTTAAACCTTA 1020
QY 1021 CACAAGGCTATTTGAAGATTTCTTAAGGCACCTGGAATTTATGATGCCACCTGTTACAGAC 1080
Db 1021 CACAAGGCTATTTGAAGATTTCTTAAGGCACCTGGAATTTATGATGCCACCTGTTACAGAC 1080
QY 1081 AATGCTAATGCAAGATGAAGCACAATGTACAGCTGGAACAGCAATCTGCACTAGAA 1140
Db 1081 AATGCTAATGCAAGATGAAGCACAATGTACAGCTGGAACAGCAATCTGCACTAGAA 1140
QY 1141 TTGTATGTAGAGGCTACAGATTTATGAAGCGGCACCTTAAGATTTGATCCCAACAA 1200
Db 1141 TTGTATGTAGAGGCTACAGATTTATGAAGCGGCACCTTAAGATTTGATCCCAACAA 1200
QY 1201 ATTGTACAAATTTGATGCTGAGAGAGATTCGGAATGTAATTCAGGAGACAGAACTAAATCT 1260
Db 1201 ATTGTACAAATTTGATGCTGAGAGAGATTCGGAATGTAATTCAGGAGACAGAACTAAATCT 1260
QY 1261 TAA 1263
Db 1261 TAA 1263
```

RESULT 3

```
US-10-364-505-2
; Sequence 2, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
; APPLICANT: Kere, Juha
; APPLICANT: Taipale, Mikko
; APPLICANT: No. US20030219787A1ola-Hemmi, Jaana
; APPLICANT: Kaminen, Nina
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0199P
; CURRENT APPLICATION NUMBER: US/10/364,505
; CURRENT FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (369)..(1628)
; FEATURE:
; OTHER INFORMATION: human DYXC1 mRNA as cDNA
US-10-364-505-2

Query Match 100.0%; Score 1263; DB 17; Length 1993;
Best Local Similarity 100.0%; Pred. No. 5.9e-253; Mismatches 0; Indels 0; Gaps 0;
Matches 1263; Conservative 0;

QY 1 ATGCTCTTTCAGGTTAGCGATTTACAGCTGGCAGCAGACGAGAACTGCGGTCTTTCTGTCT 60
Db 369 ATGCTCTTTCAGGTTAGCGATTTACAGCTGGCAGCAGACGAGAACTGCGGTCTTTCTGTCT 428
QY 61 CTGCCCCCTCAAAGGCGTGTGCTGACAGACACGAGAGTGTCTGACGGAACAACTATCTG 120
Db 429 CTGCCCCCTCAAAGGCGTGTGCTGACAGACACGAGAGTGTCTGACGGAACAACTATCTG 488
QY 121 AAGGTCACATTTTCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
Db 489 AAGGTCACATTTTCTCCATTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548
QY 181 AGCAGCAAGCAAGATTTGGAATGACACCATTTGCTTTCACCTGTATATAAAGAAAGCG 240
Db 549 AGCAGCAAGCAAGATTTGGAATGACACCATTTGCTTTCACCTGTATATAAAGAAAGCG 608
QY 241 GCCATGTGGGAGACCCCTTTCTGTGACGGTGTGACAAAGAGATGATGCAAGAAATTAGA 300
Db 609 GCCATGTGGGAGACCCCTTTCTGTGACGGTGTGACAAAGAGATGATGCAAGAAATTAGA 668
QY 301 GAAAAATCTATTTTACAAGCAACAGAGAGCAAGAAAGAGCTACAGAGCAAGAAAGCTGCA 360
Db 669 GAAAAATCTATTTTACAAGCAACAGAGAGCAAGAAAGAGCTACAGAGCAAGAAAGCTGCA 728
QY 361 CCAAGCGGAGAGATCAAAAATAGCCTAAGTGTGATGATGAAGATGAAGAAAGAG 420
Db 729 CCAAGCGGAGAGATCAAAAATAGCCTAAGTGTGATGATGAAGATGAAGAAAGAG 788
QY 421 AGGAAAAAATAGAGATATGAAGAAATGAACGGATAAAAAGCCATTAAGCATTGGAA 480
Db 789 AGGAAAAAATAGAGATATGAAGAAATGAACGGATAAAAAGCCATTAAGCATTGGAA 848
QY 481 GCCTGGAAGAAATATCAAGAAAGCTGAGAGCAAAAAAATTTAGAGAGAGAGAA 540
Db 849 GCCTGGAAGAAATATCAAGAAAGCTGAGAGCAAAAAAATTTAGAGAGAGAGAA 908
QY 541 TTATGTCACAAAGAGAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 909 TTATGTCACAAAGAGAGAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 968
QY 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAGAAATTCAGAAATATATTT 660
Db 969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAGAAATTCAGAAATATATTT 1028
QY 661 ACTGAGAGTTAAAGGAGAGACAGTATTTCTGCTCCTCGCTCTGTTGGCAGATATTAATATC 720
Db 1029 ACTGAGAGTTAAAGGAGAGACAGTATTTCTGCTCCTCGCTCTGTTGGCAGATATTAATATC 1088
QY 721 AACTTTACCCCTCGAGTATTTCCACAGCTCTTCTGTAATCAAGTACAGAGGAGAG 780
Db 1089 AACTTTACCCCTCGAGTATTTCCACAGCTCTTCTGTAATCAAGTACAGAGGAGAG 1148
QY 781 GAGTGGCTACACAAACAGCTGAGGACGAAGAGCAATGAATPACTGACATAGCTGAAT 840
Db 1149 GAGTGGCTACACAAACAGCTGAGGACGAAGAGCAATGAATPACTGACATAGCTGAAT 1208
QY 841 TCGGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 1209 TCGGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1268
QY 901 TTTGCAACGGAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAGACTAAT 960
Db 1269 TTTGCAACGGAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAGACTAAT 1328
```


Qy 961 AATAAGATGCACCTATTGTATTTGAACCGGCTGCTTGGCCACCTTAAACCTTAAACCTTA 1020
Db |||||
Qy 1329 AATAAGATGCACCTATTGTATTTGAACCGGCTGCTTGGCCACCTTAAACCTTAAACCTTA 1388
Db |||||
Qy 1021 CACAAGCTATTGAAGATTTCTTAAAGGCACTGGAATTTATTTGATGCCACCTGTTTACAGAC 1080
Db |||||
Qy 1389 CACAGGCTATTGAAGATTTCTTAAAGGCACTGGAATTTATTTGATGCCACCTGTTTACAGAC 1448
Db |||||
Qy 1081 AATGCTAATCAAGAATGAAGGCACTGTACGAGCTGGAAACAGCATTTCTGTCAACTAGAA 1140
Db |||||
Qy 1449 AATGCTAATCAAGAATGAAGGCACTGTACGAGCTGGAAACAGCATTTCTGTCAACTAGAA 1508
Db |||||
Qy 1141 TTGTATGTAGAGCCCTACAGGATTTATGAAGCGGCACTTAAGATTGATCCATCAACAAA 1200
Db |||||
Qy 1509 TTGTATGTAGAGCCCTACAGGATTTATGAAGCGGCACTTAAGATTGATCCATCAACAAA 1568
Db |||||
Qy 1201 ATTGTACAAATTTGATGCTGAGAAGATTCGGAATGTAATTCAGGAAACAGAACTAAATCT 1260
Db |||||
Qy 1569 ATTGTACAAATTTGATGCTGAGAAGATTCGGAATGTAATTCAGGAAACAGAACTAAATCT 1628
Db |||||
Qy 1261 TAA 1263
Db |||||
Qy 1629 TAA 1631

RESULT 4
US-10-681-199-2
; Sequence 2, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681.199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (369)..(1628)
; FEATURE:
; OTHER INFORMATION: human DYX1 mRNA as cDNA
US-10-681-199-2

Query Match 100.0%; Score 1263; DB 19; Length 1993;
Best Local Similarity 100.0%; Pred. No. 5.9e-253;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 60
Db |||||
Qy 369 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 428
Db |||||
Qy 61 CTGCCCCCTCAAGCGGTGTGCGTTCAGACACGCGAGCTGTTCTGCAACGAAACCTATCTG 120
Db |||||
Qy 429 CTGCCCCCTCAAGCGGTGTGCGTTCAGACACGCGAGCTGTTCTGCAACGAAACCTATCTG 488
Db |||||
Qy 121 AAGTCAACTTCTCCATTTTATTTTATTTGAGGCACTTCTTTATGCTCCCATAGACGATGAG 180
Db |||||
Qy 489 AAGTCAACTTCTCCATTTTATTTTATTTGAGGCACTTCTTTATGCTCCCATAGACGATGAG 548
Db |||||
Qy 181 AGCAGCAAGCAAGATTGGGAATTCACACCTTCTTACCTTGTATATAAAAGAGAGCG 240
Db |||||
Qy 549 AGCAGCAAGCAAGATTGGGAATTCACACCTTCTTACCTTGTATATAAAAGAGAGCG 608
Db |||||
Qy 241 GCCATGTGGGAGACCCCTTTCTGTGACCGGTTGTGCAAAAGAGATGATGCAAAAGAAATAGA 300
Db |||||
Qy 609 GCCATGTGGGAGACCCCTTTCTGTGACCGGTTGTGCAAAAGAGATGATGCAAAAGAAATAGA 668
Db |||||

RESULT 5
US-10-681-199-19
; Sequence 19, Application US/10681199
; Publication No. US20040138441A1

Qy 301 GAAAAATCTATTTTCAAGCAACAGAGAGAGCAAAAAGAGCTACAGAAGCAAAAAGCTGCA 360
Db |||||
Qy 669 GAAAAATCTATTTTCAAGCAACAGAGAGAGCAAAAAGAGCTACAGAAGCAAAAAGCTGCA 728
Db |||||
Qy 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAGAAGAG 420
Db |||||
Qy 729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAGAAGAG 788
Db |||||
Qy 421 AGCAAAAATACAGATATGAAGAAATGAAGCAATGAAGCAATGAAGCAATGAAGCAATGAAG 480
Db |||||
Qy 789 AGCAAAAATACAGATATGAAGAAATGAAGCAATGAAGCAATGAAGCAATGAAGCAATGAAG 848
Db |||||
Qy 481 GCCTGGAAGAATATCAAGAAAGCTGAGAGCAAAAAGAAATTCAGAGAGAGAGAAA 540
Db |||||
Qy 849 GCCTGGAAGAATATCAAGAAAGCTGAGAGCAAAAAGAAATTCAGAGAGAGAGAAA 908
Db |||||
Qy 541 TTATGTCAAAAGAAAGCAAAATTAAGAAAGAAAGAAAGAAAGAAATTAAGAGTCTT 600
Db |||||
Qy 909 TTATGTCAAAAGAAAGCAAAATTAAGAAAGAAAGAAAGAAAGAAATTAAGAGTCTT 968
Db |||||
Qy 601 ACTAGAAATTTGCACTAGAAATCTTCTCCAAAGGGAGAAATTCAGAGAAATATATTT 660
Db |||||
Qy 969 ACTAGAAATTTGCACTAGAAATCTTCTCCAAAGGGAGAAATTCAGAGAAATATATTT 1028
Db |||||
Qy 661 ACTGAGAAGTTAAAGGAAGACAGTATTTCTGCTCCTCGCTCTCTGTTGGCAGTATTAATC 720
Db |||||
Qy 1029 ACTGAGAAGTTAAAGGAAGACAGTATTTCTGCTCCTCGCTCTCTGTTGGCAGTATTAATC 1088
Db |||||
Qy 721 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCAAGTAGCAGAGAGGAG 780
Db |||||
Qy 1089 AACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCAAGTAGCAGAGAGGAG 1148
Db |||||
Qy 781 GAGTGCCTACACAAACAGCTGAGCGACGAGAGCAATGAATCTGACATAGCTGAAT 840
Db |||||
Qy 1149 GAGTGCCTACACAAACAGCTGAGCGACGAGAGCAATGAATCTGACATAGCTGAAT 1208
Db |||||
Qy 841 TGGGATTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 900
Db |||||
Qy 1209 TGGGATTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1268
Db |||||
Qy 901 TTTGCAACGGAAACCTATTTGGCAGCTATCAATGCATATATATTTAGCCATAAGACTAAT 960
Db |||||
Qy 1269 TTTGCAACGGAAACCTATTTGGCAGCTATCAATGCATATATATTTAGCCATAAGACTAAT 1328
Db |||||
Qy 961 AATAAGATGCACCTATTGTATTTGAACCGGCTGCTTGGCCACCTTAAACCTTAAACCTTA 1020
Db |||||
Qy 1329 AATAAGATGCACCTATTGTATTTGAACCGGCTGCTTGGCCACCTTAAACCTTAAACCTTA 1388
Db |||||
Qy 1021 CACAAGGCTATTGAAGATTTCTTCTAAGGCACTGGAATTTATGATGCCACCTGTTTACAGAC 1080
Db |||||
Qy 1389 CACAAGGCTATTGAAGATTTCTTCTAAGGCACTGGAATTTATGATGCCACCTGTTTACAGAC 1448
Db |||||
Qy 1081 AATGCTAATCAAGAATGAAGGCACTGTACGAGCTGGAAACAGCATTTCTGTCAACTAGAA 1140
Db |||||
Qy 1449 AATGCTAATCAAGAATGAAGGCACTGTACGAGCTGGAAACAGCATTTCTGTCAACTAGAA 1508
Db |||||
Qy 1141 TTGTATGTAGAGCCCTACAGGATTTATGAAGCGGCACTTAAGATTGATCCATCAACAAA 1200
Db |||||
Qy 1509 TTGTATGTAGAGCCCTACAGGATTTATGAAGCGGCACTTAAGATTGATCCATCAACAAA 1568
Db |||||
Qy 1201 ATTGTACAAATTTGATGCTGAGAAGATTCGGAATGTAATTCAGAGAAACAGAACTAAATCT 1260
Db |||||
Qy 1569 ATTGTACAAATTTGATGCTGAGAAGATTCGGAATGTAATTCAGAGAAACAGAACTAAATCT 1628
Db |||||
Qy 1261 TAA 1263
Db |||||
Qy 1629 TAA 1631

```
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pan paniscus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-19

Query Match          99.2%; Score 1253.4; DB 19; Length 1263;
Best Local Similarity 99.5%; Pred. No. 5e-251;
Matches 1257; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 60
Db 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 60

Qy 61 CTGCCCCCTCAAGGCGTGTGCTCAGACACGAGCGGTCTTCTGACCGGAAACTATCTG 120
Db 61 CTGCCCCCTCAAGGCGTGTGCTCAGACACGAGCGGTCTTCTGACCGGAAACTATCTG 120

Qy 121 AAGGTCAACTTTCCTCCATTTTATTGAGGCACTTCTTATGCTCCCATGACGATGAG 180
Db 121 AAGGTCAACTTTCCTCCATTTTATTGAGGCACTTCTTATGCTCCCATGACGATGAG 180

Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCCCTTCTATAAAAAAGAGCG 240
Db 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCCCTTCTATAAAAAAGAGCG 240

Qy 241 GCCATGTGGGAGACCTTCTCTGTGACGGGTGTGACAAAGAGATGATGCAAGAAATTAGA 300
Db 241 GCCATGTGGGAGACCTTCTCTGTGACGGGTGTGACAAAGAGATGATGCAAGAAATTAGA 300

Qy 301 GAAAAATCTATTTTACAGCAACAGAGAGAGCAAGAGAGCTACAGAGCAAAAGCTCA 360
Db 301 GAAAAATCTATTTTACAGCAACAGAGAGAGCAAGAGAGCTACAGAGCAAAAGCTCA 360

Qy 361 GCAAGCGGGAAGATCAAAAATACCACTAAGTGTGATGATGAGATGAAAGAGAGAG 420
Db 361 GCAAGCGGGAAGATCAAAAATATGCACTAAGTGTGATGATGAGATGAAAGAGAGAG 420

Qy 421 AGGAAAAAATAGAGATATGAAAGAAATGAAACGATGAAAGCACTAAAGCAATTGAA 480
Db 421 AGGAAAAAATAGAGATATGAAAGAAATGAAACGATGAAAGCACTAAAGCAATTGAA 480

Qy 481 GCCTGGAAGATATCAAGAAAGATGAGGAGCAAAAAAATTCAGAGAGAGAGAA 540
Db 481 GCCTGGAAGATATCAAGAAAGATGAGGAGCAAAAAAATTCAGAGAGAGAGAA 540

Qy 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGCAAAAAAATAAATAAAGATCTT 600
Db 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGCAAAAAAATAAATAAAGATCTT 600

Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAGGGAGAAATTCAGAAAAATATATT 660
Db 601 ACTAGAAATTTGGCATCTAGAAATCTTCTCCAAAGGGAGAAATTCAGAAAAATATATT 660

Qy 661 ACTGAGAGTTAAGGAGACAGTATTCCTGCTCCTGCTCTGTTGGCAGATTAATAATC 720
Db 661 ACTGAGAGTTAAGGAGACAGTATTCCTGCTCCTGCTCTGTTGGCAGATTAATAATC 720

Qy 721 AACTTTACCCCTCGAGTATTCACACAGCTCTTCTGTAATCAGAGTAGCAGAGAGAG 780
Db 721 AACTTTACCCCTCGAGTATTCACACAGCTCTTCTGTAATCAGAGTAGCAGAGAGAG 780
```

```
Qy 781 GAGTGGCTACACAAACAGCTGAGCAGACGAAAGCAATGAATACTGACATAGCTGAACCTT 840
Db 781 GAGTGGCTGCAACAAACAGCTGAGCAGACGAAAGCAATGAATACTGACATAGCTGAACCTT 840

Qy 841 TGCATTTTAAAGAGAGAAAGAAACCCAGAGTGGTTGAAGGATAAAGGAAACAAATG 900
Db 841 TGCATTTTAAAGAGAGAGAAAGAAACCCAGAGTGGTTGAAGGATAAAGGAAACAAATG 900

Qy 901 TTTCACACGAAAACTATTTCGACGCTATCAATGCAATATAATTTAGCCATAGACTAAT 960
Db 901 TTTCACACGAAAACTATTTCGACGCTATCAATGCAATATAATTTAGCCATAGACTAAT 960

Qy 961 AATAAGATGCCACTATTGTTTGAACCGGCTGCTGCCACCTAAACTAAAAACTT 1020
Db 961 AATAAGATGCCACTATTGTTTGAACCGGCTGCTGCCACCTAAACTAAAAACTT 1020

Qy 1021 CACAAGGCTATTGAAGATTCTTCTAAGGCATCTGGAATATTGATGCCACCTGTTACAGAC 1080
Db 1021 CACAAGGCTATTGAAGATTCTTCTAAGGCATCTGGAATATTGATGCCACCTGTTACAGAC 1080

Qy 1081 AATGCTAATCAAGAAAGAGGCAATGACGCTGGAACAGCATTCTGTCAACTAGAA 1140
Db 1081 AATGCTAATCAAGAAAGAGGCAATGACGCTGGAACAGCATTCTGTCAACTAGAA 1140

Qy 1141 TTGTATGTAGAGGCTTACAGGATTTGAAGGCGCACTTAAGATTGATCCCAACAA 1200
Db 1141 TTGTATGTAGAGGCTTACAGGATTTGAAGGCGCACTTAAGATTGATCCCAACAA 1200

Qy 1201 ATTGTACAAATTTGATGCTGAGAGATTGCGGAATGTAATTCAGAGGAACAGAACT 1260
Db 1201 ATTGTACAAATTTGATGCTGAGAGATTGCGGAATGTAATTCAGAGGAACAGAACT 1260

Qy 1261 TAA 1263
Db 1261 TAA 1263

RESULT 6
US-10-681-199-13
; Sequence 13, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-13

Query Match          99.1%; Score 1251.8; DB 19; Length 1263;
Best Local Similarity 99.4%; Pred. No. 1.1e-250;
Matches 1256; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 60
Db 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGCGGTCTTTCTGTCT 60

Qy 61 CTGCCCCCTCAAGGCGTGTGCTCAGACACGAGCGGTGTTCTGACCGGAAACTATCTG 120
Db 61 CTGCCCCCTCAAGGCGTGTGCTCAGACACGAGCGGTGTTCTGACCGGAAACTATCTG 120

Qy 121 AAGGTCAACTTTCCTCCATTTTATTGAGGCACTTCTTATGCTCCCATGACGATGAG 180
Db 121 AAGGTCAACTTTCCTCCATTTTATTGAGGCACTTCTTATGCTCCCATGACGATGAG 180
```

```
QY 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTCCACCTTGTATATAAAGAGCG 240
Db 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGCTTCCACCTTGTATATAAAGAGCG 240

QY 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300
Db 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300

QY 301 GAAAAATCTATTTTACAAGCAACAAGAGAGACAAAGAAAGCTACAGAACAAAAGCTGCA 360
Db 301 GAAAAATCTATTTTACAAGCAACAAGAGAGACAAAGAAAGCTACAGAACAAAAGCTGCA 360

QY 361 GCAAGCGGAGNATCAAAATACGCACTAAGTGTCAATGATGAAGATTGAAGAGAGAG 420
Db 361 GCAAGCGGAGNATCAAAATATGCACTAAGTGTCAATGATGAAGATTGAAGAGAGAGAG 420

QY 421 AGGAAAAATATAGAAATATGAAAGAAAAATGAAACGGATATAAAGCCACTAAAGCAATCGAA 480
Db 421 AGGAAAAATATAGAAATATGAAAGAAAAATGAAACGGATATAAAGCCACTAAAGAAATTGAA 480

QY 481 GCCTGGAAGAATATCAAAAGAAAGCTGAGAGACAAAAAATAATTCAGAGAGAGAGAAA 540
Db 481 GCCTGGAAGAATATCAAAAGAAAGCTGAGAGACAAAAAATAATTCAGAGAGAGAGAAA 540

QY 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAAAAAATAAATAATATAGAGCTTT 600
Db 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAAAAAATAAATAATATAGAGCTTT 600

QY 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 660
Db 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 660

QY 661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTGCTCTGTTGGCAGTATTAANAATC 720
Db 661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTGCTCTGTTGGCAGTATTAANAATC 720

QY 721 AACTTTACCCCTCGAGTATTCCTGCAAGCTCTTGTGAATCACAAGTAGCAGAGAGAG 780
Db 721 AACTTTACCCCTCGAGTATTCCTGCAAGCTCTTGTGAATCACAAGTAGCAGAGAGAG 780

QY 781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGCAATGAATCTGACATAGCTGAACTT 840
Db 781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGCAATGAATCTGACATAGCTGAACTT 840

QY 841 TGCATTTTAAAGAAAGAAAGAAACCCAGAAATGGTGAAGGATATAAGGAAACAAATTTG 900
Db 841 TGCATTTTAAAGAAAGAAAGAAACCCAGAAATGGTGAAGGATATAAGGAAACAAATTTG 900

QY 901 TTTGCAACGGNAACATTTTGGCAGCTATCAATGCATATTAATTTAGCCATAAGACTAAAT 960
Db 901 TTTGCAACGGNAACATTTTGGCAGCTATCAATGCATATTAATTTAGCCATAAGACTAAAT 960

QY 961 AATAAGATGCCACTATTGTTTTCGAACGGGCTCTTGCCACCTTAAACTAAAAAATCTTA 1020
Db 961 AATAAGATGCCACTATTGTTTTCGAACGGGCTCTTGCCACCTTAAACTAAAAAATCTTA 1020

QY 1021 CACAAGGCTATTGAAGATTTCTTAAGGCACTGGAATTAATTTGATGCCACCTGTTTACAGAC 1080
Db 1021 CACAAGGCTATTGAAGATTTCTTAAGGCACTGGAATTAATTTGATGCCACCTGTTTACAGAC 1080

QY 1081 AATGCTAATGCAAGAAATGAAGGCAATGTACGACGAGCAAGCAATTTCTGCAACTAGAA 1140
Db 1081 AATGCTAATGCAAGAAATGAAGGCAATGTACGACGAGCAAGCAATTTCTGCAACTAGAA 1140

QY 1141 TTGTATGTAGAGGCTTACAGATTAAGAGCGGCACTTAAGATTGATCCATCCAAACAA 1200
Db 1141 TTGTATGTAGAGGCTTACAGATTAAGAGCGGCACTTAAGATTGATCCATCCAAACAA 1200

QY 1201 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAAATTCAGGAAACAGAACTAAATCT 1260
Db 1201 ATTGTACAAATTTGATGCTGAGAGATTCGGAATGTAAATTCAGGAAACAGAACTAAATCT 1260
```

```
QY 1261 TAA 1263
Db 1261 TAA 1263

RESULT 7
US-10-681-199-15
: Sequence 15, Application US/10681199
: Publication No. US20040138441A1
: GENERAL INFORMATION:
: APPLICANT: KERE, Juha
: TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLLEXIA
: FILE REFERENCE: 0933-0214P
: CURRENT APPLICATION NUMBER: US/10/681,199
: CURRENT FILING DATE: 2003-10-09
: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 15
: LENGTH: 1263
: TYPE: DNA
: ORGANISM: Gorilla gorilla
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1260)
US-10-681-199-15

Query Match      98.9%; Score 1248.6; DB 19; Length 1263;
Best Local Similarity 99.3%; Pred. No. 5e-250;
Matches 1254; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGCCTCTTCAGCTTAGCGATTACAGCTGGCAGCAGACGAAAGACTGCGGTCTTCTCTCT 60
Db 1 ATGCCCTTCAGCTTAGCGATTACAGCTGGCAGCAGACGAAAGACTGCGGTCTTCTCTCT 60

QY 61 CTGCCCCCAAAGCGGTGTCGTCAGAGACACGACGCTGTTGACGGAATACTATCTG 120
Db 61 CTGCCCCCAAAGCGGTGTCGTCAGAGACACGACGCTGTTGACGGAATACTATCTG 120

QY 121 AAGGTCAACTTTCTCCATTTTATTTGAGGCAATTTCTTATGCTCCCATAGAGATGAG 180
Db 121 AAGGTCAACTTTCTCCATTTTATTTGAGGCAATTTCTTATGCTCCCATAGAGATGAG 180

QY 181 AGCAGCAAGCAAGAAAGATTGGGAATGACACCATTTGCTTTCACCTTGTATATAAAGAGCG 240
Db 181 AGCAGCAAGCAAGAAAGATTGGGAATGACACCATTTGCTTTCACCTTGTATATAAAGAGCG 240

QY 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300
Db 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300

QY 301 GAAAAATCTATTTTACAAGCAACAAGAGAGACAAAGAAAGCTACAGAACAAAAGCTGCA 360
Db 301 GAAAAATCTATTTTACAAGCAACAAGAGAGACAAAGAAAGCTACAGAACAAAAGCTGCA 360

QY 361 GCAAGCGGAGNATCAAAATATGCACTAAGTGTCAATGATGAAGATTGAAGAGAGAG 420
Db 361 GCAAGCGGAGNATCAAAATATGCACTAAGTGTCAATGATGAAGATTGAAGAGAGAG 420

QY 421 AGGAAAAATATAGAAATATGAAAGAAAAATGAAACGGATATAAAGCCACTAAAGCAATCGAA 480
Db 421 AGGAAAAATATAGAAATATGAAAGAAAAATGAAACGGATATAAAGCCACTAAAGAAATTGAA 480

QY 481 GCCTGGAAGAATATCAAAAGAAAGCTGAGAGACAAAAAATAATTCAGAGAGAGAGAAA 540
Db 481 GCCTGGAAGAATATCAAAAGAAAGCTGAGAGACAAAAAATAATTCAGAGAGAGAGAAA 540

QY 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAAAAAATAAATAATATAGAGCTTT 600
Db 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAAAAAATAAATAATATAGAGCTTT 600

QY 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 660
Db 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAAAATATATTT 660
```



```
QY 121 AAGTCAACTTCTCCCAATTTTATTTGAGGCAATTTCTTTATGCTCCCATAGACGATGAG 180
Db 168 AAGGTTAACTTCTCTCCCAATTTTATTTGAGCTGTTTCTATGCTCCCATAGATGCGG 227
QY 181 AGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCCACCTTGTATATAAAGCAAGCG 240
Db 228 AAGAGCAAGCAAGATTTGGGAATGACACCATTTCTTTACATTTGTATATAAAGGAGCA 287
QY 241 GCCATGTGGGAGACCTTTCTGTGACGGGTGTGTGCAAGAGAGATGATCAAGAATTAGA 300
Db 288 GTTCTGTGGATAGCTTTCTGTGCGGGTGTGTGATAAAGAGATGATGACAGAAATAGA 347
QY 301 GAAAAATCTATTTCAGACACAGAGAGACGAAAAAGAGCTACAGAGCAAAAAGCTGCA 360
Db 348 GAAAAATCTATTCTTCAAGCAGAGAGAAAGCAAGAGGCGCACAGAAAGCAAAAGCTGTT 407
QY 361 GCAAGGCGGGAAGATCAAAAATPAGCACTTAAGTGTGATGATGAAGATTTGAAGAAGAG 420
Db 408 GCCAAGCAGAGAGACAGAGATACCACTTAGCGAGATGATGAAGATTTGAAGAAGAGAG 467
QY 421 AGGAAAAATATAGAAATATGAAAGAAATGAACGGATAAAGCCACTAAAGCAATTGAA 480
Db 468 AGGAAAAATCTCGAAGATCTGAAGAAATGAACGGAAAGAGCAAACTAGCGAATTAGA 527
QY 481 GCCTGGAAGAAATATCAAGAAAGAGCTGAGGAGCAAAAAGAAATTCAGAGAGAGAGAA 540
Db 528 GCGTGGAAAGAAATGTCAAAAGAAAGCTGACGCAAAAAGAGAGTCCAGAGGAAGAGAA- 586
QY 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 587 --ACGCTCGAGGGAAGCAAGCT---GAAGAGACCAAAAGCTCTAAAACCTCGGGGTTG 641
QY 601 ACTAGAAATTTGGCACTTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAAATATTT 660
Db 642 CCCGGAAGGCCCACTCGCTCCCAAGAGGGAGGAATTTGGGAAAAACATATTT 701
QY 661 ACTGAGAAATTAAGGAAGACGATATCTGCTCTCTGCTCTGTTGGCAGTATTAAGATC 720
Db 702 CCTGAGAAATTAAGGAAGACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 761
QY 721 AACTTTACCTCCGAGTATCCCAACAGCTCTTGTGATCATCAGTACGAGAGAGAGAG 780
Db 762 AGCTTTACCTCCGAGTATCCCAACAGCTCTTGTGGAATCCCAAGTCCGAGAGAGAGAG 821
QY 781 GAGTGGCTACACAAAGAGCTGAGGCAAGAGCAATGAATATCTGACATAGCTGAACCT 840
Db 822 GAGTGGCTGCATNAACAGAGAGAGCAAGAGAGGCAATGAGCACTGACCTTCTGAGTTC 881
QY 841 TGGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 882 TTTGACTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 941
QY 901 TTTGCAAGGAAAGCAATTTTGGCAGCTATCAATGATATATTTAGCCATTAAGCAATAAT 960
Db 942 TTTGCAACAGAGAAACTATTGCGCAGCGTTGATGATATATTTAGCCATTAAGCAATAAT 1001
QY 961 AATAGATGCGCACTATTGTTTGAACCGGCTGCTGCGCACTTAAACTTAAAGAACTTTA 1020
Db 1002 TGTAGATGCCATTTATGTTGATTTGATCGGCTGCTGCGCACTTAAACTTAAAGAACTTTA 1061
QY 1021 CACAAGGCTATTTGAAGATTTCTTTAAGGCACTGGAATTTATGATGCCACCTGTTTACAG 1080
Db 1062 CACAAGGCTATGAGGACTCTTTTAAAGCACTAGAGATTTATGATGCCACCTGTTTACAG 1121
QY 1081 AATGCTAATGCAAGATGAGGCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1122 AATGCAATGCAAGATGAGGCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1181
QY 1141 TTGATGATGAGAGGCTACAGGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1182 TTGATGATGAGAGGCTTGAAGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1241
```

```
QY 1201 ATTGTACAAATTTGATGCTGACAGAGATTCGGAAATGTAATTCAGGACAGAACTAAATCT 1260
Db 1242 GTTGTACAGAAAGATGACAGAGAGATTCGGAATATATTAATTCAGGGACGCGACTGAAGTCT 1301

RESULT 11
US-10-681-199-4
; Sequence 4, Application US/10681199
; Publication No. US2004013841A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1697
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(1307)
; FEATURE:
; OTHER INFORMATION: murine DYX1 mRNA as cDNA
US-10-681-199-4

Query Match 66.6%; Score 840.8; DB 19; Length 1697;
Best Local Similarity 80.3%; Pred. No. 4.7e-165;
Matches 1012; Conservative 0; Mismatches 242; Indels 6; Gaps 2;

QY 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAAGACTGCGGTCTTCTGTCT 60
Db 48 ATGCAGTGCAGTGAGCGAATTACGCTGGCAGCAGACGCGGCCACGATCTTCTGTGCG 107
QY 61 CTGCCCCCTAAAGCGGTGCTGCTGACAGACAGCGAGCTGTTCTGACCGGAAACTATCTG 120
Db 108 CTGCTCTGCGGGGCGTCTGCTGCGGATCTGACGTAATCTGTTGGGGAAGTTACCTG 167
QY 121 AAGTCAACTTCTCCCAATTTTATTTGAGGCAATTTCTTTATGCTCCCATAGACGATGAG 180
Db 168 AAGTTAACTTCTCTCCATTTTATTTGAGCTGTTTCTATGCTCCCATAGATGATGGG 227
QY 181 AGCAGCAAGCAAGATTTGGGAATGACACCATTTGTTTCACTTTGTATATAAAGAGCG 240
Db 228 AAGAGCAAGCAAGATTTGGGAATGACACGATTTCTTTTCACTTGTATAAAGAGGAGCA 287
QY 241 GCCATGTGGGAGACCTTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAATTAGA 300
Db 288 GTTCTGTGGATAGCTTTCTGTGCGGGTGTGATAAAGAGATGATGACAGAAATAAGA 347
QY 301 GAAAAATCTATTTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 348 GAAAAATCTATTCTTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407
QY 361 GCAAGGCGGGAAGATCAAAAATAGCCACTAAGTGTGATGATGAAGATTTGAAGAAGAGAG 420
Db 408 GCCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467
QY 421 AGGAAAAATATAGAAATATGAAAGAAATGAACGGATAAAGCCACTAAAGCAATTGAA 480
Db 468 AGGAAAAATCTCGAAGATCTGAAAGAAATGAACGGAAAAAGGCAACTAGCGAATTAGA 527
QY 481 GCCTGGAAGAGATCAAGAGAAAGAGCTGAGGAGCAAAAAGAAATTCAGAGAGAGAGAGAA 540
Db 528 GCGTGGAAAGAAATGTCAAAAGAAAGCTGACGCAAAAAGAGAGTCCAGAGGAAGAGAGAA- 586
QY 541 TTATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 587 --ACGCTCGAGGGAAGCAAGCT---GAAGAGACCAAAAGCTCTAAAACCTCGGGGTTG 641
QY 601 ACTAGAAATTTGGCACTTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAAATATTT 660
```



```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (347)..(347)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-30001

Query Match      31.8%; Score 401.6; DB 18; Length 464;
Best Local Similarity 98.3%; Pred. No. 8.7e-74;
Matches 404; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGAGCAAGACTGCGTCTTCTGCT 60
DB 49 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGAGCAAGACTGCGTCTTCTGCT 108
QY 61 CTGCCCCCTCAAGGGCGTGTGCGTCTAGAGACACGACGCTGTTCTGACCGAAAACTATCTG 120
DB 109 CTGCCCCCTCAAGGGCGTGTGCGTCTAGAGACACGACGCTGTTCTGACCGAAAACTATCTG 168
QY 121 AGGTCAACTTTCCTCCATTTTATTTATTTAGGCAATTTCTTATGCTCCCATAGACGATGAG 180
DB 169 AGGTCAACTTTCCTCCATTTTATTTATTTAGGCAATTTCTTATGCTCCCATAGACGATGAG 228
QY 181 AGCAGCAAGCAAGATGGGAATGACACCATTTGCTTACCTCTGTATATAAAAAAGACG 240
DB 229 AGCAGCAAGCAAGATGGGAATGACACCATTTGCTTACCTCTGTATATAAAAAAGACG 288
QY 241 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGTGACAAAGAGATGATGCAAGAAATTAGA 300
DB 289 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGTGACAAAGAGATGATGCAAGAAATTANA 348
QY 301 GAAAAATCTATTTTCAAGCAACAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 360
DB 349 GAAAAATCTATTTTCAAGCAACAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 408
QY 361 GCAAGCGGAGAGTCAAAATACCACTAGTGTGATGATGATGAAATTTGAA 411
DB 409 GCAAGCGGAGAGTCAAAATACCACTAGTGTGATGATGATGAGGGAGAA 459

RESULT 14
US-09-918-995-22675
; Sequence 22675, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22675
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10303

Query Match      29.3%; Score 369.8; DB 10; Length 488;
Best Local Similarity 99.5%; Pred. No. 3.7e-67;
Matches 371; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 35 AGACGAAGACTGCGGTCTTCTGCTCTGCTCCCTCAAGGGGTGTGCGTCTAGAGACAGG 94
DB 28 AGACGAAGACTGCGGTCTTCTGCTCTGCTCCCTCAAGGGGTGTGCGTCTAGAGACAGG 87
QY 95 ACGTGTCTGACGGAAGAACTATCTGAAGGTCAACTTTCTCCATTTTATTTAGGGCAT 154
DB 88 ACGTGTCTGACGGAAGAACTATCTGAAGGTCAACTTTCTCCATTTTATTTAGGGCAT 147

; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (347)..(347)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-30001

Query Match      31.8%; Score 401.6; DB 18; Length 464;
Best Local Similarity 98.3%; Pred. No. 8.7e-74;
Matches 404; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGAGCAAGACTGCGTCTTCTGCT 60
DB 49 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGAGCAAGACTGCGTCTTCTGCT 108
QY 61 CTGCCCCCTCAAGGGCGTGTGCGTCTAGAGACACGACGCTGTTCTGACCGAAAACTATCTG 120
DB 109 CTGCCCCCTCAAGGGCGTGTGCGTCTAGAGACACGACGCTGTTCTGACCGAAAACTATCTG 168
QY 121 AGGTCAACTTTCCTCCATTTTATTTATTTAGGCAATTTCTTATGCTCCCATAGACGATGAG 180
DB 169 AGGTCAACTTTCCTCCATTTTATTTATTTAGGCAATTTCTTATGCTCCCATAGACGATGAG 228
QY 181 AGCAGCAAGCAAGATGGGAATGACACCATTTGCTTACCTCTGTATATAAAAAAGACG 240
DB 229 AGCAGCAAGCAAGATGGGAATGACACCATTTGCTTACCTCTGTATATAAAAAAGACG 288
QY 241 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGTGACAAAGAGATGATGCAAGAAATTAGA 300
DB 289 GCCATGTGGGAGACCCCTTCTGTGACGGGTGTGTGACAAAGAGATGATGCAAGAAATTANA 348
QY 301 GAAAAATCTATTTTCAAGCAACAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 360
DB 349 GAAAAATCTATTTTCAAGCAACAGAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGCA 408
QY 361 GCAAGCGGAGAGTCAAAATACCACTAGTGTGATGATGATGAAATTTGAA 411
DB 409 GCAAGCGGAGAGTCAAAATACCACTAGTGTGATGATGATGAGGGAGAA 459

RESULT 15
US-09-918-995-10303
; Sequence 10303, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10303
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10303

Query Match      26.6%; Score 336; DB 10; Length 458;
Best Local Similarity 91.1%; Pred. No. 4e-60;
Matches 368; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 639 GAGAAATTCAGAAATATATTTTACTGAGAACTTAAAGGAAGACAGTATTCCTGCTCTCG 698
DB 56 GAGAAAGACTGCAATATATTTTACTGAGACGTTAAAGGAAGACAGTATTTACTGCACTCT 115
QY 699 CTCTGTTGGCAGTATTTAAATCAACTTTACCCCTCGAGTATTTCCCAACAGCTCTTCGTA 758
DB 116 GTGTGTCATGCGCTA-CGCCACCAACCCAGTCGAGAAATTTCCAATCAGTGTCTTCCTGA 174
QY 759 ATCACAAGTAGCAGAGAGGAGGAGTGGCTACACAAACAAAGCTGAGGACCAAGAGCAAT 818
DB 175 ATTCAAGTAACACAGAGGAGGAGTGGCTACACAAACAGCTGAGGACCAAGAGCAAT 234
QY 819 GAATCTGACATAGCTGAACCTTTGCGATTTTAAAGGAAGAAAGAAACCCAGAAATGGTT 878
DB 235 GAATCTGACATAGCTGAACCTTTGCGATTTTAAAGGAAGAAAGAAACCCAGAAATGGTT 294
QY 879 GAAGGATAAGGAAACAAATTTGTTTGGCAACCGGAAAACTATTTGGCAGCTATCAATGATA 938
DB 295 GAAGGATAAGGAAACAAATTTGTTTGGCAACCGGAAAACTATTTGGCAGCTATCAATGATA 354
QY 939 TAATTTAGCCATAGACTTAATATATAGTCCCACTATTTGATTTGAACCGGCTGCTTG 998
DB 355 TAATTTAGCCATAGACTTAATATATAGTCCCACTATTTGATTTGAACCGGCTGCTTG 414
```

Qy 999 CCACCTAAACTAAAACTTACACAGGCTATTGAAGATTCTT 1042
Ob 415 CCACCTAAACTAAAACTTACACAGGCTATTGAAGATTCTT 458

Search completed: July 8, 2005, 07:35:57
Job time : 1946 secs

Result	Query	Score	Length	ID	Description
No.	Match		DB		

1	1263	100.0	1263	17	US-10-364-505-1	Sequence 1, Appl
2	1263	100.0	1263	19	US-10-681-199-1	Sequence 1, Appl
3	1263	100.0	1993	17	US-10-364-505-2	Sequence 2, Appl
4	1263	100.0	1993	19	US-10-681-199-2	Sequence 2, Appl
5	1002	79.3	1263	19	US-10-681-199-19	Sequence 19, Appl
6	998	79.0	1559	17	US-10-108-260A-575	Sequence 575, App
7	960	76.0	1263	19	US-10-681-199-13	Sequence 13, Appl

```

RESULT 1
US-10-364-505-1
; Sequence 1, Application US/10364505-1
; Publication No. US20030219787A1
; GENERAL INFORMATION:
; APPLICANT: Kere, Juha
; APPLICANT: Taipale, Mikko
; APPLICANT: No. US20030219787A1ola
; APPLICANT: Kaminen, Nina
; TITLE OF INVENTION: NOVEL HUMAN GE
; FILE REFERENCE: 0933-0199P
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequenc
US-10-364-505-1

```

Query Match	100.0%;	Score 1263;	DB 17;	Length 1263;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1263;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCGCTCTTCAGGTTAGCGATTACAGCTGCGCAGCAGACGGAAGACTTCGGGTCTTCTGCT	60	
Db	1	ATGCGCTCTTCAGGTTAGCGATTACAGCTGCGCAGCAGACGGAAGACTTCGGGTCTTCTGCT	60	

ALIGNMENTS

```

RESULT 1
US-10-364-505-1
; Sequence 1, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
; APPLICANT: Kere, Juhna
; APPLICANT: Taipale, Mikko
; APPLICANT: No. US20030219787A1ola-Hemmi, Jaana
; APPLICANT: Kaminen, Nina
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLLEXIA
; FILE REFERENCE: 0933-0199P
; CURRENT APPLICATION NUMBER: US/10/364,505
; CURRENT FILING DATE: 2003-02-12

```

```

; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding
US-10-364-505-1

```

Qy 61 CTGCCCCCAAAGCGGTGGGTGAGAGACGACGAGCGTGTCTGACGGAAACATATCTG 120
Db |||||
Qy 61 CTGCCCCCAAAGCGGTGGGTGAGAGACGACGAGCGTGTCTGACGGAAACATATCTG 120
Db |||||
Qy 121 AAGGTCAACTTCTCCATTTTATTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
Db |||||
Qy 121 AAGGTCAACTTCTCCATTTTATTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
Db |||||
Qy 181 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGCTTTACCTTGTATATAAAGAAAGCG 240
Db |||||
Qy 181 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGCTTTACCTTGTATATAAAGAAAGCG 240
Db |||||
Qy 241 GCCATGTGGAGACCCCTTTCTGTACCGGTGTTGACAAAGAGATGACAAAGAAATTAGA 300
Db |||||
Qy 241 GCCATGTGGAGACCCCTTTCTGTACCGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300
Db |||||
Qy 301 GAAAAATCTATTTTACAGACACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db |||||
Qy 301 GAAAAATCTATTTTACAGACACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db |||||
Qy 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 420
Db |||||
Qy 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 420
Db |||||
Qy 421 AGGAAAAATAGAAATATGAAAGAAATGAACGGATTAAGCCACTAAAGCTTGGAA 480
Db |||||
Qy 421 AGGAAAAATAGAAATATGAAAGAAATGAACGGATTAAGCCACTAAAGCTTGGAA 480
Db |||||
Qy 481 GCCTGGAAAGAAATATCAAGAAAGCTGAGGACAAAGAAATTTACAGAGAAAGAGAAA 540
Db |||||
Qy 481 GCCTGGAAAGAAATATCAAGAAAGCTGAGGACAAAGAAATTTACAGAGAAAGAGAAA 540
Db |||||
Qy 541 TTATGTCAAAGAAAGCAAAATTAAGAAAGAAAGAAAGAAATTAATAATTAAGAGTCTT 600
Db |||||
Qy 541 TTATGTCAAAGAAAGCAAAATTAAGAAAGAAAGAAAGAAATTAATAATTAAGAGTCTT 600
Db |||||
Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAATATATTT 660
Db |||||
Qy 601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAGGGAGAAATTCAGAAATATATTT 660
Db |||||
Qy 661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCTGCTGTTGGCAGTATTAAATC 720
Db |||||
Qy 661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCTGCTGTTGGCAGTATTAAATC 720
Db |||||
Qy 721 AACTTTACCCCTCGAGTATTCCTGCTCTGCTGATCAACAAGTACAGAGAGAGGAG 780
Db |||||
Qy 721 AACTTTACCCCTCGAGTATTCCTGCTCTGCTGATCAACAAGTACAGAGAGAGGAG 780
Db |||||
Qy 781 GAGTGGCTACAAAACAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAACCTT 840
Db |||||
Qy 781 GAGTGGCTACAAAACAGCTGAGGACGAGAGCAATGAATCTGACATAGCTGAACCTT 840
Db |||||
Qy 841 TGGCATTTAAAGGAAGAAAGAACCCAGATGGTTGAGAGTAAAGGAACAAATTTG 900
Db |||||
Qy 841 TGGCATTTAAAGGAAGAAAGAACCCAGATGGTTGAGAGTAAAGGAACAAATTTG 900
Db |||||
Qy 901 TTTGCAACGGAAAACTATTTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960
Db |||||
Qy 901 TTTGCAACGGAAAACTATTTTGGCAGCTATCAATGCATATATTTAGCCATAAGACTAAAT 960
Db |||||
Qy 961 AATAAGTGCACATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1020
Db |||||
Qy 961 AATAAGTGCACATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1020
Db |||||
Qy 1021 CACAAGCTATTGAAGATTTCTTCAAGGACCTGGAATTTTATGATGCCACCTGTTTACAGAC 1080
Db |||||
Qy 1021 CACAAGCTATTGAAGATTTCTTCAAGGACCTGGAATTTTATGATGCCACCTGTTTACAGAC 1080
Db |||||
Qy 1081 AATGCTAATGCAAGAAATGAAGGACATGTACGAGCTGGAAACAGCATTTCTGTCAACTAGAA 1140
Db |||||
Qy 1081 AATGCTAATGCAAGAAATGAAGGACATGTACGAGCTGGAAACAGCATTTCTGTCAACTAGAA 1140
Db |||||

Qy 1141 TTGTTATGTAGAACGCTTACAGGATTATGAAGCGCACTTAAGATTGTATCCATCCAAACAA 1200
Db |||||
Qy 1141 TTGTTATGTAGAACGCTTACAGGATTATGAAGCGCACTTAAGATTGTATCCATCCAAACAA 1200
Db |||||
Qy 1201 ATTGTACAAATTGATGCTGAGAAAGATTCCGAATGTAAATCAAGGAACAGAACTAAAAATCT 1260
Db |||||
Qy 1201 ATTGTACAAATTGATGCTGAGAAAGATTCCGAATGTAAATCAAGGAACAGAACTAAAAATCT 1260
Db |||||
Qy 1261 TAA 1263
Db |||||
Qy 1261 TAA 1263
Db |||||
RESULT 2
US-10-681-199-1
; Sequence 1, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequence for human DYX1 (cDNA)
US-10-681-199-1
Query Match 100.0%; Score 1263; DB 19; Length 1263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGGGTCTTTCTCTCT 60
Db 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGGGTCTTTCTCTCT 60
Qy 61 CTGCCCCCAAAGCGGTGGGTGAGAGACGACGAGCGTGTCTGACGGAAACATATCTG 120
Db 61 CTGCCCCCAAAGCGGTGGGTGAGAGACGACGAGCGTGTCTGACGGAAACATATCTG 120
Qy 121 AAGTCAACTTCTCCATTTTATTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
Db 121 AAGTCAACTTCTCCATTTTATTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
Qy 181 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGCTTTACCTTGTATATAAAGAAAGCG 240
Db 181 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGCTTTACCTTGTATATAAAGAAAGCG 240
Qy 241 GCCATGTGGAGACCCCTTTCTGTACCGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300
Db 241 GCCATGTGGAGACCCCTTTCTGTACCGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 300
Qy 301 GAAAAATCTATTTTACAGACACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Db 301 GAAAAATCTATTTTACAGACACAGAGAGAGCAAAAGAGCTACAGAAAGCAAAAGCTGCA 360
Qy 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 420
Db 361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCTATGATGAAGATTGAAGAAAGAG 420
Qy 421 AGGAAAAATAGAAATATGAAAGAAATGAACGGATTAAGCCACTAAAGCTTGGAA 480
Db 421 AGGAAAAATAGAAATATGAAAGAAATGAACGGATTAAGCCACTAAAGCTTGGAA 480
Qy 481 GCCTGGAAAGAAATATCAAGAAAGCTGAGGACAAAGAAATTTACAGAGAAAGAGAAA 540
Db 481 GCCTGGAAAGAAATATCAAGAAAGCTGAGGACAAAGAAATTTACAGAGAAAGAGAAA 540
Db |||||

Qy	961	AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTAA	1020
Db	1329	AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTAA	1388
Qy	1021	CACAAGCTATTGAAGATTTCTTCTAAGGCACCTGGAATTTATGATGCCACCTGTTACAGAC	1080
Db	1389	CACAAGCTATTGAAGATTTCTTCTAAGGCACCTGGAATTTATGATGCCACCTGTTACAGAC	1448
Qy	1081	AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTTCTGCTCAACTAGAA	1140
Db	1449	AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTTCTGCTCAACTAGAA	1508
Qy	1141	TTGTATGTAGAAGSCCTACAGGATTTATGAAGCGGCACCTTAAGATTGATCCATCCAAACAAA	1200
Db	1509	TTGTATGTAGAAGSCCTACAGGATTTATGAAGCGGCACCTTAAGATTGATCCATCCAAACAAA	1568
Qy	1201	ATTGTACAATTTGATGCTGTAGAGATTCGGAAATGTAATTCACAGACAGAACTAAAATCT	1260
Db	1569	ATTGTACAATTTGATGCTGTAGAGATTCGGAAATGTAATTCACAGACAGAACTAAAATCT	1628
Qy	1261	TAA 1263	
Db	1629	TAA 1631	

```

RESULT 4
US-10-681-199-2
; Sequence 2, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 1993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (369)..(1628)
; FEATURE:
; OTHER INFORMATION: human DYX1C1 mRNA as cDNA
US-10-681-199-2

```

Qy	301	GA	AAAA	TCTAT	TTTT	CA	AG	CA	GA	AG	GAG	CA	AA	AA	GA	AG	CT	C	A	GA	AG	CA	AA	AA	AG	CT	G	CA	360				
																												728					
Db	669	GA	AAAA	TCTAT	TTTT	CA	AG	CA	GA	AG	GAG	CA	AA	AA	GA	AG	CT	C	A	GA	AG	CA	AA	AA	AG	CT	G	CA	728				
Qy	361	GC	AA	CG	GG	GA	AG	CA	GT	CA	AA	AA	AT	AC	CA	AT	AC	CA	AT	AC	CA	AT	AC	CA	AT	AC	CA	AT	420				
Db	729	GC	AA	CG	GG	GA	AG	CA	GT	CA	AA	AA	AT	AC	CA	AT	AC	CA	AT	AC	CA	AT	AC	CA	AT	AC	CA	AT	788				
Qy	421	AG	GA	AAAA	AT	TA	GA	GA	TAT	GA	AA	GA	AAAA	TG	AA	CG	GA	T	A	A	A	A	A	A	A	A	A	A	480				
Db	789	AG	GA	AAAA	AT	TA	GA	GA	TAT	GA	AA	GA	AAAA	TG	AA	CG	GA	T	A	A	A	A	A	A	A	A	A	A	848				
Qy	481	GC	CT	GA	GA	GA	GA	AT	TA	CA	AG	AA	AA	GT	CG	AG	CA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	540				
Db	849	GC	CT	GA	GA	GA	GA	AT	TA	CA	AG	AA	AA	GT	CG	AG	CA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	908				
Qy	541	TT	AT	GT	C	MA	AA	GA	AA	AG	CA	AA	TT	AA	AA	GA	AG	AA	GA	AA	AA	AA	AA	AA	AA	AA	AA	AA	600				
Db	909	TT	AT	GT	C	MA	AA	GA	AA	AG	CA	AA	TT	AA	AA	GA	AG	AA	GA	AA	AA	AA	AA	AA	AA	AA	AA	AA	968				
Qy	601	ACT	GA	AA	AT	TT	GC	AT	CT	AG	AA	AT	CT	TT	GC	T	C	AA	AA	GG	GA	AA	AT	T	C	AG	AA	AT	TAT	660			
Db	969	ACT	GA	AA	AT	TT	GC	AT	CT	AG	AA	AT	CT	TT	GC	T	C	AA	AA	GG	GA	AA	AT	T	C	AG	AA	AT	TAT	1028			
Qy	661	ACT	GA	GA	AG	TT	AA	GG	AA	GA	CA	GT	AT	CT	CG	CT	CG	CT	CG	CT	CG	CT	CG	CT	CG	CT	CG	CT	720				
Db	1029	ACT	GA	GA	AG	TT	AA	GG	AA	GA	CA	GT	AT	CT	CG	CT	CG	CT	CG	CT	CG	CT	CG	CT	CG	CT	CG	CT	1088				
Qy	721	AA	CT	TT	AC	CC	CT	CG	AG	TAT	T	CC	CA	AG	CT	CT	TT	CG	TA	AT	C	A	CA	AG	T	AG	CA	AG	780				
Db	1089	AA	CT	TT	AC	CC	CT	CG	AG	TAT	T	CC	CA	AG	CT	CT	TT	CG	TA	AT	C	A	CA	AG	T	AG	CA	AG	1148				
Qy	781	G	AG	T	G	G	C	T	A	C	A	AA	A	C	AA	G	T	G	A	G	C	A	A	A	G	C	A	A	A	840			
Db	1149	G	AG	T	G	G	C	T	A	C	A	AA	A	C	AA	G	T	G	A	G	C	A	A	A	G	C	A	A	A	1208			
Qy	841	T	G	C	GA	TT	T	A	A	A	G	A	GA	AA	AA	GA	CC	C	A	GA	AT	G	T	T	G	A	A	G	A	900			
Db	1209	T	G	C	GA	TT	T	A	A	A	G	A	GA	AA	AA	GA	CC	C	A	GA	AT	G	T	T	G	A	A	G	A	1268			
Qy	901	TT	T	G	C	A	CG	GA	AA	CT	AT	TT	T	G	C	A	G	CT	AT	CA	AT	GC	AT	TA	AT	TT	T	AG	CA	TA	AA	AT	960
Db	1369	TT	T	G																													

RESULT 5
US-10-681-199-19
; Sequence 19, Application US/10681199
; Publication No. US20040138441A1

```
; GENERAL INFORMATION:
; APPLICANT: KRE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pan paniscus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-19

Query Match          79.3%; Score 1002; DB 19; Length 1263;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 CTTCAAGGCTGTGCGTCAGACACGACGCGTGTCTGACGGAAAACTATCTGAAGTTC 126
DB 7 CTTCAAGGCTGTGCGTCAGACACGACGCGTGTCTGACGGAAAACTATCTGAAGTTC 66
QY 67 CTCAAGGCTGTGCGTCAGACACGACGCGTGTCTGACGGAAAACTATCTGAAGTTC 126
DB 67 CTCAAGGCTGTGCGTCAGACACGACGCGTGTCTGACGGAAAACTATCTGAAGTTC 126
QY 127 AACTTTCCTCCATTTTATTTTTCAGGCAATTTCTTATGCTCCCATAGAGCGAGGAGC 186
DB 127 AACTTTCCTCCATTTTATTTTTCAGGCAATTTCTTATGCTCCCATAGAGCGAGGAGC 186
QY 187 AAAGCAAGATTTGGCAATGACACCAATTTCTTCACTTGTATATAAAAGAGCGCCATG 246
DB 187 AAAGCAAGATTTGGCAATGACACCAATTTCTTCACTTGTATATAAAAGAGCGCCATG 246
QY 247 TGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATGACAAAGAAATAGAGAAAA 306
DB 247 TGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATGACAAAGAAATAGAGAAAA 306
QY 307 TCTATTTTACAGCACAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGACAGCAAG 366
DB 307 TCTATTTTACAGCACAGAGAGCAAAAGAGCTACAGAGCAAAAGCTGACAGCAAG 366
QY 367 CGGGAAGATCAAAAATACGCATTAAGTGTCTATGATGAAGATTGAAGAAGAGAGAGAAA 426
DB 367 CGGGAAGATCAAAAATATGCACTAAGTGTCTATGATGAAGATTGAAGAAGAGAGAGAAA 426
QY 427 AAAATAGAGATATGAAGAAATGAACGGATATAAAGCCACTTAAAGCAATTTGAAGCCTGG 486
DB 427 AAAATAGAGATATGAAGAAATGAACGGATATAAAGCCACTTAAAGCAATTTGAAGCCTGG 486
QY 487 AAGGATATCAAGAAAGCTGAGGACGCAAAAAAATTCAGAGAGAGAGAAATATGT 546
DB 487 AAGGATATCAAGAAAGCTGAGGACGCAAAAAAATTCAGAGAGAGAGAAATATGT 546
QY 547 CAAAAAGAAAGCAATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
DB 547 CAAAAAGAAAGCAATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
QY 607 AATTGGCATCTAGAAATCTTGTCTCAAAAGGGAGAGAAATTCAGAGAGAGAGAGAGAGAG 666
DB 607 AATTGGCATCTAGAAATCTTGTCTCAAAAGGGAGAGAAATTCAGAGAGAGAGAGAGAGAG 666
QY 667 AAGTTAAAGAGAGACAGTATTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
DB 667 AAGTTAAAGAGAGACAGTATTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
QY 727 ACCCTCGAGTATTCCTCAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
DB 727 ACCCTCGAGTATTCCTCAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
```

```
QY 787 CTACACAAAGCTGAGCGACGAGCAAGCAATGAATACTGACATAGCTGAACCTTTGCGAT 846
DB 787 CTGCAACAAAGCTGAGCGACGAGCAAGCAATGAATACTGACATAGCTGAACCTTTGCGAT 846
QY 847 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
DB 847 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
QY 907 ACGGAAAACTATTTGGCAGCTATCAATGCATATATAATTTAGCCATTAAGACTAAATAAG 966
DB 907 ACAGAAAACTATTTGGCAGCTATCAATGCATATATAATTTAGCCATTAAGACTAAATAAG 966
QY 967 ATGCCACTATTTGATTTGAAACCGGGCTGCTTGGCCACTTAAACTTAAACTTACACAG 1026
DB 967 ATGCCACTATTTGATTTGAAACCGGGCTGCTTGGCCACTTAAACTTAAACTTACACAG 1026
QY 1027 GCTATTGAAGATTTCTTAAAGCACTGGAATTTATTGATGCCACCTGTTACAGACAATGCT 1086
DB 1027 GCTATTGAAGATTTCTTAAAGCACTGGAATTTATTGATGCCACCTGTTACAGACAATGCT 1086
QY 1087 AATGCAAGATTAAGGCAATGTACGACGTGGAAACAGCATTTCTGTTCAACTAGAAATGTA 1146
DB 1087 AATGCAAGATTAAGGCAATGTACGACGTGGAAACAGCATTTCTGTTCAACTAGAAATGTA 1146
QY 1147 GTAGAGCCCTACAGGATTTAAGAGCGGCACTTAAGATTTGATCCATCCAAACAAATTTGA 1206
DB 1147 GTAGAGCCCTACAGGATTTAAGAGCGGCACTTAAGATTTGATCCATCCAAACAAATTTGA 1206
QY 1207 CAAATTGATGCTGAGAGATTTGGAATTTAAATTTCAAGGAAACAGAACTTAAATCTTAA 1263
DB 1207 CAAATTGATGCTGAGAGATTTGGAATTTAAATTTCAAGGAAACAGAACTTAAATCTTAA 1263

RESULT 6
US-10-108-260A-575
; Sequence 575, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 575
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-575

Query Match          79.0%; Score 998; DB 17; Length 1559;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGACGAGCTGCGGTCTTCTGTCT 60
DB 43 ATGCTCTTTCAGGTTAGCGATTACAGCTGGCAGACGAGAGCTGCGGTCTTCTGTCT 102
QY 61 CTGCCCTTCAAGGCGGTGCTGCTGAGACACGACGCTGTTCTGACCGGAAACTATCTG 120
DB 103 CTGCCCTTCAAGGCGGTGCTGCTGAGACACGACGCTGTTCTGACCGGAAACTATCTG 162
QY 121 AAGTCAACTTCTCCATTTTATTTGAGGCAATTTTATGCTCCCATAGACGATGAG 180
DB 163 AAGTCAACTTCTCCATTTTATTTGAGGCAATTTTATGCTCCCATAGACGATGAG 222
QY 181 AGCAGCAAGCAAGATTTGGGAATGACACCAATTTCTTCACTTTGTATATAAAAGAGCG 240
DB 223 AGCAGCAAGCAAGATTTGGGAATGACACCAATTTCTTCACTTTGTATATAAAAGAGCG 282
QY 241 GCCATGTGGAGAGACCTTTCTGTCGGGTGTTGACAAAGAGATGATGCAAGAAATTTAGA 300
DB 241 GCCATGTGGAGAGACCTTTCTGTCGGGTGTTGACAAAGAGATGATGCAAGAAATTTAGA 300
```


283 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTAGA 342
Qy
301 GAAAAATCTATTTTACAAGCAACAAGAGAGAGCAAAAAGAGCTACAGAAGCAAAAAGCTGCA 360
Db
343 GAAAAATCTATTTTACAAGCACACAGAGAGAGCAAAAAGAGCTACAGAAGCAAAAAGCTGCA 402
Qy
361 GCAAAGCGGAAGATCAAAAATACGCACTAAGTGTCAATGATGAAGATTGAAGAAGAGAG 420
Db
403 GCAAAGCGGAAGATCAAAAATACGCACTAAGTGTCAATGATGAAGATTGAAGAAGAGAG 462
Qy
421 AGGAAAAAATAGAAAGATATGAAGAATAATGAACGGATATAAAGCCACTAAAGCAATTGGAA 480
Db
463 AGGAAAAAATAGAAAGATATGAAGAATAATGAACGGATATAAAGCCACTAAAGCAATTGGAA 522
Qy
481 GCCTGGAAAGATATCAAAAGAAAGCTGAGAGCAAAAAAATTTCAAGAGAGAGAGAAA 540
Db
523 GCCTGGAAAGATATCAAAAGAAAGCTGAGAGCAAAAAAATTTCAAGAGAGAGAGAAA 582
Qy
541 TTAATGTCAAAAGAAAGCAAAATTAAGAAGGAAGAAATAAATAAATAAAGAGTCTT 600
Db
583 TTAATGTCAAAAGAAAGCAAAATTAAGAAGGAAGAAATAAATAAATAAAGAGTCTT 642
Qy
601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAAATATATT 660
Db
643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAAATATATT 702
Qy
661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCGCTGTTGGCAGTATTAATAATC 720
Db
703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCGCTGTTGGCAGTATTAATAATC 762
Qy
721 AACTTTACCCCTCGAGTATTCCTGCTCCTCGCTCGCTGTTGGCAGTATTAATAATC 780
Db
763 AACTTTACCCCTCGAGTATTCCTGCTCCTCGCTCGCTGTTGGCAGTATTAATAATC 822
Qy
781 GAGTGGCTACACAAACAAAGCTGAGGACGAGCAATGATGATGATGATGATGATGATGATGAT 840
Db
823 GAGTGGCTACACAAACAAAGCTGAGGACGAGCAATGATGATGATGATGATGATGATGATGAT 882
Qy
841 TGCATTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 900
Db
883 TGCATTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 942
Qy
901 TTTGCAAGCGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 960
Db
943 TTTGCAAGCGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1002
Qy
961 AATAAGATGCCATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1020
Db
1003 AATAAGATGCCATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1062

RESULT 7
US-10-681-199-13
; Sequence 13, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681.199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (1)..(1260)
US-10-681-199-13
Query Match 76.0%; Score 960; DB 19; Length 1263;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1210; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 49 GTCTTTCTGTCTGTGCGCCCTCAAAAGCGTGTGCGTCTAGAGACACGAGCGTGTCTGCAACG 108
Db 49 GTCTTTCTGTCTGTGCGCCCTCAAAAGCGTGTGCGTCTAGAGACACGAGCGTGTCTGCAACG 108
Qy 109 GAAAACTATCTGAAGTCAACTTTCTCTCCATTTTATTTATTTGAGGCATTTCTTTATGCTCCC 168
Db 109 GAAAACTATCTGAAGTCAACTTTCTCTCCATTTTATTTATTTGAGGCATTTCTTTATGCTCCC 168
Qy 169 ATAGACGATGAGAGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCACTTGTAT 228
Db 169 ATAGACGATGAGAGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTCACTTGTAT 228
Qy 229 AAAAAAGAGCGGCCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288
Db 229 AAAAAAGAGCGGCCCATGTGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288
Qy 289 CAAAGAAATTAGAGAAAAATCTATTTTACAAGCAACAAGAGAGCAAAAGAGCTTACAGAA 348
Db 289 CAAAGAAATTAGAGAAAAATCTATTTTACAAGCAACAAGAGAGCAAAAGAGCTTACAGAA 348
Qy 349 GCAAAAGCTGAGCAAAAGCGGGAAGATCAAAAATACCCACTAAGTGTGATGATGATGATG 408
Db 349 GCAAAAGCTGAGCAAAAGCGGGAAGATCAAAAATATGCACTAAGTGTGATGATGATGATG 408
Qy 409 GAAGAGAGAGAGGAAAAAATAGAGATATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGA 468
Db 409 GAAGAGAGAGAGGAAAAAATAGAGATATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGA 468
Qy 469 AAAGCATTTGAAGCCTTGGAAAGAAATATCAAAAGAAAGCTGAGGAGCAAAAAAATTTTCA 528
Db 469 AAAGCATTTGAAGCCTTGGAAAGAAATATCAAAAGAAAGCTGAGGAGCAAAAAAATTTTCA 528
Qy 529 AGAGAGAGAAATTAATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAG 588
Db 529 AGAGAGAGAAATTAATGTCAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAG 588
Qy 589 TATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGTCTCCAAAAGGAGAGAAATCA 648
Db 589 TATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGTCTCCAAAAGGAGAGAAATCA 648
Qy 649 GAAAAATATTTTACTGAGAAAGTTAAAGAGAGACAGTATTTCTGCTCCTGCTCTGTTGGC 708
Db 649 GAAAAATATTTTACTGAGAAAGTTAAAGAGAGACAGTATTTCTGCTCCTGCTCTGTTGGC 708
Qy 709 AGTATTAATAATCAACTTTTACCCCTCGAGTATTTCCCAAACAGCTCTTCTGTAATCAAGTA 768
Db 709 AGTATTAATAATCAACTTTTACCCCTCGAGTATTTCCCAAACAGCTCTTCTGTAATCAAGTA 768
Qy 769 GCAGAGAGAGGAGTGGCTACACAAACAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
Db 769 GCAGAGAGAGGAGTGGCTACACAAACAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
Qy 829 ATAGCTCAACTTTTTCGCAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
Db 829 ATAGCTCAACTTTTTCGCAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
Qy 889 GGAAGCAAAATTTGTTGCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
Db 889 GGAAGCAAAATTTGTTGCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
Qy 949 ATAGAGCTTAATAATAGATGCCCACTATTTGATTTGAAACCGGGCTGCTTGGCAGCTAAAA 1008
Db 949 ATAGAGCTTAATAATAGATGCCCACTATTTGATTTGAAACCGGGCTGCTTGGCAGCTAAAA 1008
Qy 1009 CTAAAAAATTTTACCAAGGCTATTGAAGATTTCTTAAAGGCACTTGAATTTGATGCA 1068
Db 1009 CTAAAAAATTTTACCAAGGCTATTGAAGATTTCTTAAAGGCACTTGAATTTGATGCA 1068

Db 1009 CTAAABAACTTACACAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCA 1068
Qy 1069 CCTGTTACAGACAATGCTAATGCAAGATGAAGGACACATGTACGACGTGGGAACAGCATTC 1128
Db 1069 CCTGTTACAGACAATGCTAATGCAAGATGAAGGACACATGTACGACGTGGGAACAGCATTC 1128
Qy 1129 TGTCAACTAGAAATCTATGTAGAGCGCTTACAGGATTAAGCGGCACTTTAAGATTGAT 1188
Db 1129 TGTCAACTAGAAATCTATGTAGAGCGCTTACAGGATTAAGCGGCACTTTAAGATTGAT 1188
Qy 1189 CCATCCAAACAAATTTGTAACAATTCATGCTGAGAAAGATTCGGAATGTAATTCAGGAACA 1248
Db 1189 CCATCCAAACAAATTTGTAACAATTCATGCTGAGAAAGATTCGGAATGTAATTCAGGAACA 1248
Qy 1249 GAACATAAAATCTTAA 1263
Db 1249 GAACATAAAATCTTAA 1263

RESULT 8

US-10-681-199-15
; Sequence 15, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-15

Query Match 68.0%; Score 859; DB 19; Length 1263;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 48 GGTCTTTCTGCTCTGCCCCCTCAAGGGCGTGTGCGTCAGACACGCGACGTGTTCTGCAC 107
Db 48 GGTCTTTCTGCTCTGCCCCCTCAAGGGCGTGTGCGTCAGACACGCGACGTGTTCTGCAC 107
Qy 108 GGAACACTCTGAAGGTCACACTTTCCTCCATTTTATTTGAGGCACTTCTTTATGCTCC 167
Db 108 GGAACACTCTGAAGGTCACACTTTCCTCCATTTTATTTTATTTGAGGCACTTCTTTATGCTCC 167
Qy 168 CATGACCATGAGACGACGACAAAGATGGAATGGAATGACACCATGCTTCTCACTCTGTA 227
Db 168 CATGACCATGAGACGACGACAAAGATGGAATGGAATGACACCATGCTTCTCACTCTGTA 227
Qy 228 TAAAAAGAGCGGCATGTGGGAGCCCTTTCTGTGACGGGTGTGACAAAGAGATGAT 287
Db 228 TAAAAAGAGCGGCATGTGGGAGCCCTTTCTGTGACGGGTGTGACAAAGAGATGAT 287
Qy 288 GCAAGAAATAGAGAAAATCTATTTTACAGCACACAGAGAGACAAAGAGCTACAGA 347
Db 288 GCAAGAAATAGAGAAAATCTATTTTACAGCACACAGAGAGACAAAGAGCTACAGA 347
Qy 348 AGCAAAAGCTCAGCAAAAGCGGAGATGACAAAATACGCACCTAAAGTGTGATGATGAT 407
Db 348 AGCAAAAGCTCAGCAAAAGCGGAGATGACAAAATATGCAATATGCAATGTCATGATGAT 407
Qy 408 TGAAG 467
Db 408 TGAAG 467
Qy 468 TAAAGCATTTGGAAGCCTCGAAGAAATATCAAGAAAGAGCTGAGGAGCAAAAAAATTC 527

Db 468 TAAAGAAATTGAGAGCCTGGAAAGAAATATCAAGAAAGAGCTGAGGAGCAAAAAAATTC 527
Qy 528 GAGAGAGAGAGAAATTTATGTCAAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAG 587
Db 528 GAGAGAGAGAGAAATTTATGTCAAAAAGAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAG 587
Qy 588 ATATAAGAGCTTACTAGAGAAATTTGGCATCTAGAGAAATCTTGGTCCAAAAGGGAGAAATTC 647
Db 588 ATATAAGAGCTTACTAGAGAAATTTGGCATCTAGAGAAATCTTGGTCCAAAAGGGAGAAATTC 647
Qy 648 AGAAATATATTTACTGAGAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
Db 648 AGAAATATATTTACTGAGAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
Qy 708 CAGTATTAAAAATCAACTTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCAAGT 767
Db 708 CAGTATTAAAAATCAACTTTTACCCCTCGAGTATTTCCCAACAGCTCTTCTGTAATCAAGT 767
Qy 768 AGCAG 827
Db 768 AGCAG 827
Qy 828 CATAGCTGAACCTTTGCGAATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 887
Db 828 CATAGCTGAACCTTTGCGAATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 887
Qy 888 AGGAAACAAATTTGTTGCAACGAGAAACTATTTGGCAGCTATCAATGCAATATAATTTAGC 947
Db 888 AGGAAACAAATTTGTTGCAACGAGAAACTATTTGGCAGCTATCAATGCAATATAATTTAGC 947
Qy 948 CATAGACTAAATTAATTAAGATGCCACTATTTGTAATTTGAAACCGGCTGCTGGCACCTAAA 1007
Db 948 CATAGACTAAATTAATTAAGATGCCACTATTTGTAATTTGAAACCGGCTGCTGGCACCTAAA 1007
Qy 1008 ACTAAAAAATTTACAGAGGCTATTTGAAGATTTCTTAAGGCACTTGAATTTATGATGCC 1067
Db 1008 ACTAAAAAATTTACAGAGGCTATTTGAAGATTTCTTAAGGCACTTGAATTTATGATGCC 1067
Qy 1068 ACCTGTTACAGCAATGCTAATGCAAGAAATGAAGGCACTATGACACGTGGAAACAGCAT 1127
Db 1068 ACCTGTTACAGCAATGCTAATGCAAGAAATGAAGGCACTATGACACGTGGAAACAGCAT 1127
Qy 1128 CTGTCAACTAGAAATGTAATGTAAGAGGCTTACAGGATTAAGCGGCACTTAAGATTTGA 1187
Db 1128 CTGTCAACTAGAAATGTAATGTAAGAGGCTTACAGGATTAAGCGGCACTTAAGATTTGA 1187
Qy 1188 TCCATCCAAACAAATTTGTAAGAAATTTGATGCTGAGAGATTCGGAATGTAATTCAGGAAC 1247
Db 1188 TCCATCCAAACAAATTTGTAAGAAATTTGATGCTGAGAGATTCGGAATGTAATTCAGGAAC 1247
Qy 1248 AGAACTAAAAATCTTAA 1263
Db 1248 AGAACTAAAAATCTTAA 1263

RESULT 9

US-10-681-199-17
; Sequence 17, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pongo pygmaeus
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (1)...(1260)
US-10-681-199-17

Query Match      55.1%; Score 696; DB 19; Length 1263;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1246; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 7 CTTGAGTTAGCGATTACAGCTGGCAGCAGACGAGAGCTGGGTCTTTCTGTCTCTGCC 66
Db 7 CTTGAGTTAGCGATTACAGCTGGCAGCAGACGAGAGCTGGGTCTTTCTGTCTCTGCC 66

Qy 67 CTCAAGGCGTGTCCGTGAGAGACGAGCGTGTCTGACGAGAACTATCTGAGGTC 126
Db 67 CTCAAGGCGTGTCCGTGAGAGACGAGCGTGTCTGACGAGAACTATCTGAGGTC 126

Qy 127 AACTTTCTCCATTTTATTTAGGCAATTTCTTATGCTCCCATAGACGATGAGAGCAGC 186
Db 127 AACTTTCTCCATTTTATTTAGGCAATTTCTTATGCTCCCATAGACGATGAGAGCAGC 186

Qy 187 AAAGCAAGATTGGGAATGACACCATTTCTTCACTTGTATATAAAGAGCGGCCATG 246
Db 187 AAAGCAAGATTGGGAATGACACCATTTCTTCACTTGTATATAAAGAGCGGCCATG 246

Qy 247 TGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTTAGAGAAA 306
Db 247 TGGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTTAGAGAAA 306

Qy 307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGACTACAGAAAGCTGCAAGCAAAAG 366
Db 307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGACTACAGAAAGCTGCAAGCAAAAG 366

Qy 367 CGGGAAGATCAAAATACGCACTAAGTGTCTGATGATGAGATTTGAAGAAAGAGAGAAA 426
Db 367 CGGGAAGATCAAAATATGCACTAAGTGTCTGATGATGAGATTTGAAGAAAGAGAGAAA 426

Qy 427 AAAATAGAGATGAAAGAAATGAAAGGATGAAAGGCTAAAGCAATTTGAGAGCTGG 486
Db 427 AAAATAGAGATGAAAGAAATGAAAGGATGAAAGGCTAAAGCAATTTGAGAGCTGG 486

Qy 487 AAAGAAATCAAGAAAGCTGAGAGCAAAAGAAATTTAGAGAGAGAGAAATTTATGT 546
Db 487 AAAGAAATCAAGAAAGCTGAGAGCAAAAGAAATTTAGAGAGAGAGAAATTTATGT 546

Qy 547 CAAAAAGAAAGCAAAATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
Db 547 CAAAAAGAAAGCAAAATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606

Qy 607 AATTGGCCTAGAAATCTTGCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
Db 607 AATTGGCCTAGAAATCTTGCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666

Qy 667 AAGTTAAAGAGAGAGAGATTTCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 726
Db 667 AAGTTAAAGAGAGAGAGATTTCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 726

Qy 727 ACCCTCGAGATTTCCCAAGAGCTCTTCTGTAATCAAGAGTACAGAGAGAGAGAGAGAG 786
Db 727 ACCCTCGAGATTTCCCAAGAGCTCTTCTGTAATCAAGAGTACAGAGAGAGAGAGAGAG 786

Qy 787 CTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
Db 787 CTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846

Qy 847 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
Db 847 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906

Qy 907 ACAGAAAGATTTTGGAGAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAAG 966
Db 907 ACAGAAAGATTTTGGAGAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAAG 966

Qy 967 ATGCCACTATTGTTGAAACGGGCTGTTGCCACCTTAAACTTAAAGAACTTACACAAAG 1026
```

```
Db 967 ATGCCACTATTGTTGAAACGGGCTGTTGCCACCTTAAACTTAAAGAACTTACACAAAG 1026

Qy 1027 GCTATTGAGAGATTTCTTAAAGGCACTGGAATTAATGATGCACTGTGTACAGACAATGCT 1086
Db 1027 GCTATTGAGAGATTTCTTAAAGGCACTGGAATTAATGATGCACTGTGTACAGACAATGCT 1086

Qy 1087 AATGCAAGATGAGGACACATGTACGACGTGGAAACAGCATTTCTGTCACTAGAAATGCTAT 1146
Db 1087 AATGCAAGATGAGGACACATGTACGACGTGGAAACAGCATTTCTGTCACTAGAAATGCTAT 1146

Qy 1147 GTAGAAGCCCTACAGGATTAAGCGGCACTTAAAGATTGATCCATCCAAACAAATTTGTA 1206
Db 1147 GTAGAAGCCCTACAGGATTAAGCGGCACTTAAAGATTGATCCATCCAAACAAATTTGTA 1206

Qy 1207 CAAATGATGCTGAGAAAGATTCGGAATGTAATTTCAAGGAAACAGAACTTAAATCTTAA 1263
Db 1207 CAAATGATGCTGAGAAAGATTCGGAATGTAATTTCAAGGAAACAGAACTTAAATCTTAA 1263

RESULT 10
US-09-918-995-22675
; Sequence 22675, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22675
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-22675

Query Match      25.8%; Score 326; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 1.7e-151;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 GCGTCAGAGACACGAGCGTGTCTGACGGAAACTATCTGAAGTCACTTCTCCAT 139
Db 73 GCGTCAGAGACACGAGCGTGTCTGACGGAAACTATCTGAAGTCACTTCTCCAT 132

Qy 140 TTTTATTTGAGGCAATTTCTTATGCTCCCATAGACGATGAGAGAGAGAGAGAGAGAG 199
Db 133 TTTTATTTGAGGCAATTTCTTATGCTCCCATAGACGATGAGAGAGAGAGAGAGAGAG 192

Qy 200 GGAATGACACATTTCTTCACTTTGTATAAAGAGAGCGGCCATGTGGAGACCCCTTT 259
Db 193 GGAATGACACATTTCTTCACTTTGTATAAAGAGAGCGGCCATGTGGAGACCCCTTT 252

Qy 260 CTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTTAGAGAAATCTATTTTACAAG 319
Db 253 CTGTGACGGGTGTTGACAAAGAGATGATGCAAGAAATTTAGAGAAATCTATTTTACAAG 312

Qy 320 CACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
Db 313 CACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372

Qy 380 AATACGCACTAAGTGTCTATGATGAAG 405
Db 373 AATACGCACTAAGTGTCTATGATGAAG 398

RESULT 11
US-10-425-872/c
; Sequence 872, Application US/10240425
```

Publication No. US20040033502A1

GENERAL INFORMATION:
APPLICANT: Williams, Amanda
APPLICANT: Bolland, Joseph F.
APPLICANT: Lord, Reginald V.
APPLICANT: Alvarez, Chris
APPLICANT: Wetzel, Jon C.
APPLICANT: Scherff, Uwe
APPLICANT: Vockley, Joseph G.
TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
FILE REFERENCE: 44921-5026
CURRENT APPLICATION NUMBER: US/10/240,425
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: PCT/US01/09847
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 60/193,446
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 1588
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 872
LENGTH: 325
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20040033502A1 AI783611
US-10-240-425-872

Query Match 25.7%; Score 325; DB 18; Length 325;
Best Local Similarity 100.0%; Pred. No. 5.2e-151;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 TTCAGAGAGAGAGAAATATGTCAAAAGAAAAGCAAAATTAAGAGAGAGAGAAAAA 583
DB 325 TTCAGAGAGAGAGAAATATGTCAAAAGAAAAGCAAAATTAAGAGAGAGAGAAAAA 266
QY 584 TAAATATAAGAGTCTTACTAGAAATTTGGCAATCTAGAAATCTGCTCCAAAGGAGAA 643
DB 265 TAAATATAAGAGTCTTACTAGAAATTTGGCAATCTAGAAATCTGCTCCAAAGGAGAA 206
QY 644 ATTCAGAAAATATTTACTGAGAAAGTTAAAGGAGAGAGATTTCTGCTCTCTGCTG 703
DB 205 ATTCAGAAAATATTTACTGAGAAAGTTAAAGGAGAGAGATTTCTGCTCTCTGCTG 146
QY 704 TTGGCAGTATTAATAATCAACTTTACCCCTCGAGTATTCCTCAACAGCTCTTCTGTAATCAC 763
DB 145 TTGGCAGTATTAATAATCAACTTTACCCCTCGAGTATTCCTCAACAGCTCTTCTGTAATCAC 86
QY 764 AAGTAGCAGAGAGAGAGGAGTGCTTACACAAACAGCTGAGGCAGAGAGCAATGAATA 823
DB 85 AAGTAGCAGAGAGAGAGGAGTGCTTACACAAACAGCTGAGGCAGAGAGCAATGAATA 26
QY 824 CTGACATAGCTCAACTTTTGGCAATTT 848
DB 25 CTGACATAGCTCAACTTTTGGCAATTT 1

RESULT 12

US-09-918-995-10303
Sequence 10303, Application US/09918995
Publication No. US20030073623A1

GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10303
LENGTH: 458

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(458)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10303

Query Match 21.4%; Score 270; DB 10; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.4e-123;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 773 AAGAGGAGAGTGGCTACACAAACAGCTTGAGGACGAGAGCAATGATCTGACATAG 832
DB 189 AAGAGGAGAGTGGCTACACAAACAGCTTGAGGACGAGAGCAATGATCTGACATAG 248
QY 833 CTGAACCTTTGGCATTTAAAGAGAGAGAAAGAACCCAGAAATGGTTGAAGGATAAGGAA 892
DB 249 CTGAACCTTTGGCATTTAAAGAGAGAGAAAGAACCCAGAAATGGTTGAAGGATAAGGAA 308
QY 893 ACAAATTTGTCACCGAAAACTATTTGGCAGCTATCAATGCATATATTTAGGCATAA 952
DB 309 ACAAATTTGTCACCGAAAACTATTTGGCAGCTATCAATGCATATATTTAGGCATAA 368
QY 953 GACTAAATAATAAGATGCCATATTTGTAATTTGAACCGGCTGCTGCCACCTAAACTAA 1012
DB 369 GACTAAATAATAAGATGCCATATTTGTAATTTGAACCGGCTGCTGCCACCTAAACTAA 428
QY 1013 AAACTTACACAGGCTATTGAAGATCTT 1042
DB 429 AAACTTACACAGGCTATTGAAGATCTT 458

RESULT 13

US-10-956-157-3114/c
Sequence 3114, Application US/10956157
Publication No. US20050118625A1

GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Mounts, William
TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
FILE REFERENCE: 031896-043000 (AM 101081)
CURRENT APPLICATION NUMBER: US/10/956,157
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 319805
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3114
LENGTH: 715
TYPE: DNA
ORGANISM: Homo sapiens
US-10-956-157-3114

Query Match 20.0%; Score 253; DB 21; Length 715;
Best Local Similarity 100.0%; Pred. No. 4.2e-115;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 638 GGAGAAATTCAGAAAAATATTTACTGAGAAAGTTAAAGGAGAGAGATTTCTGCTCTC 697
DB 430 GGAGAAATTCAGAAAAATATTTACTGAGAAAGTTAAAGGAGAGAGATTTCTGCTCTC 371
QY 698 GCTCTGTTGGCAGTATTAATAATCACTTTACCCCTCGAGTATTCCTCAACAGCTCTTCTG 757
DB 370 GCTCTGTTGGCAGTATTAATAATCACTTTACCCCTCGAGTATTCCTCAACAGCTCTTCTG 311
QY 758 AATCAAGTAGCAGAGAGAGGAGTGCTTACACAAACAGCTGAGCAGAGAGAGAA 817
DB 310 AATCAAGTAGCAGAGAGAGGAGTGCTTACACAAACAGCTGAGCAGAGAGAGAA 251
QY 818 TGAATCTGACATAGCTGAATTTTGGCAATTTTAAAGAGAGAGAAAGAAACCCAGAAATG 877
DB 250 TGAATCTGACATAGCTGAATTTTGGCAATTTTAAAGAGAGAGAAAGAAACCCAGAAATG 191

Qy 878 TGAAGGATAAAGG 890
Db 190 TGAAGGATAAAGG 178

RESULT 14

US-10-956-157-8349
; Sequence 8349, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; NUMBER OF SEQ ID NOS: 2004-10-04
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8349
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-8349

Query Match 20.0%; Score 253; DB 21; Length 715;
Best Local Similarity 100.0%; Pred. No. 4.2e-115;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 638 GGAGAAATTCAGAAAATATATTACTGAGAAAGTTAAAGGAGACAGATTCTCTCTC 697
Db 286 GGAGAAATTCAGAAAATATATTACTGAGAAAGTTAAAGGAGACAGATTCTCTCTC 345
Qy 698 GCTCTGTTGGCAGTATTAAATCAACTTTACCCCTCGAGTATCCCAACAGCTCTTCGTG 757
Db 346 GCTCTGTTGGCAGTATTAAATCAACTTTACCCCTCGAGTATCCCAACAGCTCTTCGTG 405
Qy 758 AATCACAAGTAGCAG 817
Db 406 AATCACAAGTAGCAG 465
Qy 818 TGAATCTGACATAGCTGAACCTTTGCGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAG 877
Db 466 TGAATCTGACATAGCTGAACCTTTGCGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAG 525
Qy 878 TGAAGGATAAAGG 890
Db 526 TGAAGGATAAAGG 538

RESULT 15

US-10-242-535A-30001
; Sequence 30001, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30001
; LENGTH: 464

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (263)..(263)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (321)..(321)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (347)..(347)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-30001

Query Match 16.9%; Score 214; DB 17; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.1e-95;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGGGCTCTTCTCTCT 60
Db 49 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAGACTGGGCTCTTCTCTCT 108
Qy 61 CTGCCCTCAAAAGCGGTGCGTCAGAGACACGAGACGTCTTCGACGAGAAACTATCTG 120
Db 109 CTGCCCTCAAAAGCGGTGCGTCAGAGACACGAGACGTCTTCGACGAGAAACTATCTG 168
Qy 121 AAGTCAACTTTCTCTCCATTTTATTTAGGCAATTTCTTTATGCTCCCATAGACGATGAG 180
Db 169 AAGTCAACTTTCTCTCCATTTTATTTAGGCAATTTCTTTATGCTCCCATAGACGATGAG 228
Qy 181 AGCAGCAAGCAAAAGATTGGGAATGACACCATTTG 214
Db 229 AGCAGCAAGCAAAAGATTGGGAATGACACCATTTG 262

Search completed: July 8, 2005, 16:12:46
Job time : 865 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 06:20:46 ; Search time 9619 Seconds
(without alignments)
4997.939 Million cell updates/sec

Title: US-10-681-199-1

Perfect score: 1263
Sequence: 1 atgccttcaggttagcga.....gaacagaactaaatcttaa 1263

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	838	66.3	1600	3	BC017392	BC017392 Homo sapi
2	583	46.2	844	6	CD107587	CD107587 AGENCOURT
3	462	36.6	535	7	CN478982	CN478982 UI-CF-FNO
4	446	35.3	1131	5	BQ217312	BQ217312 AGENCOURT
5	422	33.4	462	7	CN429788	CN429788 170006000
6	410	32.5	788	2	BE564350	BE564350 601343161
7	405	32.1	823	5	BU567908	BU567908 AGENCOURT
8	378	29.9	792	4	BG771796	BG771796 602720472
9	374	29.6	933	6	CD358543	CD358543 AGENCOURT
10	364	28.8	573	2	BE972748	BE972748 601652170
11	342	27.1	772	6	CA422987	CA422987 UI-H-FLO
12	340	26.9	468	6	CB297042	CB297042 12B22058
13	338	26.8	417	1	A1024221	A1024221 ov81e02.s
14	325	25.7	325	1	A1783611	A1783611 tz99c01.x
15	323	25.6	651	5	BU852580	BU852580 AGENCOURT
16	320	25.3	701	4	BG482634	BG482634 602502783
17	313	24.8	1183	2	BF207765	BF207765 601861861
18	290	23.0	665	4	BG540324	BG540324 602568825
19	275	21.8	559	5	BU607404	BU607404 UI-CF-FNO
20	274	21.7	625	2	BF216970	BF216970 601884034
21	251	19.9	645	5	BM984145	BM984145 UI-CF-DUI
22	250	19.8	270	2	BF372375	BF372375 PM4-FT002
23	244	19.3	744	4	BG192162	BG192162 RST11269
24	241	19.1	267	1	A1073572	A1073572 ov45g11.x

C 25	218	17.3	583	6	CB160627	CB160627 K-EST02020
C 26	216	17.1	382	1	AA724419	AA724419 ah91f07.s
C 27	215	17.0	775	4	BG206612	BG206612 RST26063
C 28	212	16.8	523	2	BE178005	BE178005 RC3-HT060
C 29	197	15.6	251	1	AA317003	AA317003 EST18868
C 30	182	14.4	454	2	BE463906	BE463906 hv18f08.x
C 31	175	13.9	988	4	BI517373	BI517373 603041624
C 32	150	11.9	752	4	BG183140	BG183140 RST2158.A
C 33	148	11.7	940	4	BG220651	BG220651 RST40439
C 34	146	11.6	855	2	BF248143	BF248143 601859338
C 35	105	8.3	467	1	AI674107	AI674107 wd18c04.x
C 36	102	8.1	500	4	BI517981	BI517981 603041624
C 37	85	6.7	459	5	EX095509	EX095509 BX095509
C 38	78	6.2	402	2	BE178258	BE178258 RC3-HT060
C 39	78	6.2	683	5	BP460416	BP460416 BP460416
C 40	70	5.5	112	4	EM820152	EM820152 K-EST0088
C 41	64	5.1	440	1	AI360851	AI360851 qx99h12.x
C 42	61	4.8	793	7	CO738958	CO738958 SILR04C21
C 43	60	4.8	654	5	BM971229	BM971229 UI-CF-DUI
C 44	57	4.5	515	5	BU738973	BU738973 UI-E-EJ0-
C 45	51	4.0	285	8	AQ067717	AQ067717 HS_2239_B

ALIGNMENTS

RESULT 1
BC017392
LOCUS Homo sapiens, Similar to RIKEN cDNA 1700010124 gene, clone
DEFINITION IMAGE:4081622, mRNA.
ACCESSION BC017392
VERSION BC017392.1 GI:19263480
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1600)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-4590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
<http://www.systembiology.org>
contact: amadane@systembiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IPAL Plate: 32 Row: k Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction
This clone has the following problem: frame shifted.

FEATURES
source Location/Qualifiers
1..1600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4081622"
/tissue_type="Bladder, carcinoma"
/clone_lib="NIH_MGC_53"
/lab_host="DH10B"

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13293 row: n column: 13
 High quality sequence stop: 388.
 Location/Qualifiers
 FEATURES
 source
 1. .1131
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6047052"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN
 Query Match 35.3%; Score 446; DB 5; Length 1131;
 Best Local Similarity 100.0%; Pred. No. 9.1e-217; Indels 0; Gaps 0;
 Matches 446; Conservative 0; Mismatches 0;

Qy 42 GACTGCGGTCTTTCTGTCCTGCCCTCAAGGCGTGTGCATGACACACGACGCTGT 101
 Db 1 GACTGCGGTCTTTCTGTCCTGCCCTCAAGGCGTGTGCATGACACACGACGCTGT 60
 Qy 102 CTGCACGGAAACTATCTGAGGTCACCTTCTCCATTTTATTTGAGGCATTTCTTTA 161
 Db 61 CTGCACGGAAACTATCTGAGGTCACCTTCTCCATTTTATTTGAGGCATTTCTTTA 120
 Qy 162 TGCTCCCATAGACATGACAGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTTCAC 221
 Db 121 TGCTCCCATAGACATGACAGCAGCAAGCAAGATTTGGGAATGACACCATTTGCTTTCAC 180
 Qy 222 CTTGTATAAAAAAGAACGGCCATGTGGAGACCCCTTCTGTGACGGGTGTGACAAGA 281
 Db 181 CTTGTATAAAAAAGAACGGCCATGTGGAGACCCCTTCTGTGACGGGTGTGACAAGA 240
 Qy 282 GATGATCGAAGATTAGAGAAATCTTTTACAGCACACAGAGAGCAAGAAAGC 341
 Db 241 GATGATCGAAGATTAGAGAAATCTTTTACAGCACACAGAGAGCAAGAAAGC 300
 Qy 342 TACAGAGCAAAAGCTGCAGCAAGCGGAAGATCAAAAATACGCATTAAGTGTCTATGAT 401
 Db 301 TACAGAGCAAAAGCTGCAGCAAGCGGAAGATCAAAAATACGCATTAAGTGTCTATGAT 360
 Qy 402 GAAGATTGAAGAAGAGAGGAAATAATAGAGATATGAAGAAATGAACGGATAAA 461
 Db 361 GAAGATTGAAGAAGAGAGGAAATAATAGAGATATGAAGAAATGAACGGATAAA 420
 Qy 462 AGCCACTAAAGCATTTGAGCCTGGA 487
 Db 421 AGCCACTAAAGCATTTGAGCCTGGA 446

RESULT 5
 CN429788
 LOCUS CN429788 462 bp mRNA linear EST 16-MAY-2004
 DEFINITION 17000600026401 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
 CN429788
 VERSION CN429788.1 GI:47417382
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 462)
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
 Lebkowski, J. and Stanton, L.W.
 Transcriptome characterization elucidates signaling networks that
 control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 462 Std Error: 0.00.

FEATURES
 source

1. .462
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, DMSO-treated H9 cell
 line"
 /clone_lib="GRN_PREHEP"
 /note="Oligo dT primed, full-length enriched cDNA library
 from DMSO-treated hES cell line H9 (p22) maintained in
 feeder-free conditions"

ORIGIN

Query Match 33.4%; Score 422; DB 7; Length 462;
 Best Local Similarity 100.0%; Pred. No. 1.7e-204; Indels 0; Gaps 0;
 Matches 422; Conservative 0; Mismatches 0;

Qy 1 ATGCTCTTTCAGGTAGCGATTACAGCTGGCAGCAGACGAAAGACTGCGGTCTTTCTGTCT 60
 Db 41 ATGCTCTTTCAGGTAGCGATTACAGCTGGCAGCAGACGAAAGACTGCGGTCTTTCTGTCT 100
 Qy 61 CTGCCCTCAAAGCGGTGCGTCAGAGACACGAGCGTGTTCGACGGAAACTATCTG 120
 Db 101 CTGCCCTCAAAGCGGTGCGTCAGAGACACGAGCGTGTTCGACGGAAACTATCTG 160
 Qy 121 AAGGTCAACTTTCCTCCATTTTATTTGAGGCATTTCTTTATGTCCTCCATAGAGATGAG 180
 Db 161 AAGGTCAACTTTCCTCCATTTTATTTGAGGCATTTCTTTATGTCCTCCATAGAGATGAG 220
 Qy 181 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGTCTTCACTTGTATAAAAAAGAACGC 240
 Db 221 AGCAGCAAAAGCAAGATTGGGAATGACACCATTTGTCTTCACTTGTATAAAAAAGAACGC 280
 Qy 241 GCATGTGGAGACCCCTTCTGTGACGGGTGTGACAAAGAGATGATGCAAAAGATTAGA 300
 Db 281 GCCATGTGGAGACCCCTTCTGTGACGGGTGTGACAAAGAGATGATGCAAAAGATTAGA 340
 Qy 301 GAAAAATCTATTTTACAGCACACAGAGAGCAAAAGAACTACAGAAAGCTGCA 360
 Db 341 GAAAAATCTATTTTACAGCACACAGAGAGCAAAAGAACTACAGAAAGCTGCA 400
 Qy 361 GCAAAGCGGGAAGATCAAAAATACGCCAATAGTGTCTATGATGAAGATTGAAGAGAAAGAG 420
 Db 401 GCAAAGCGGGAAGATCAAAAATACGCCAATAGTGTCTATGATGAAGATTGAAGAGAAAGAG 460
 Qy 421 AG 422
 Db 461 AG 462

RESULT 6
 BE564350
 LOCUS BE564350 788 bp mRNA linear EST 15-AUG-2000
 DEFINITION 601343161F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3685335 5',
 mRNA sequence.
 BE564350
 BE564350
 BE564350.1 GI:9808070
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM373 row: 1 column: 16
High quality sequence stop: 556.
Location/Qualifiers

FEATURES

1. 788
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3685335"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 53"
/note="Organ: bladder; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcttcggcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 32.5%; Score 410; DB 2; Length 788;
Best Local Similarity 100.0%; Pred. No. 2.5e-198; Mismatches 0; Indels 0; Gaps 0;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGCAGAGACTGCGGTCTTTCTGTCT 60
Db 23 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGCAGAGACTGCGGTCTTTCTGTCT 82
Qy 61 CTGCCCCCTCAAGGCGTGTGCTCAGACACGACGCTTCTGCGGAAACTATCTG 120
Db 83 CTGCCCCCTCAAGGCGTGTGCTCAGACACGACGCTTCTGCGGAAACTATCTG 142
Qy 121 AAGGTCAACTTTCTCCATTTTATTTCAGGCAATTCCTTATGCTCCCATAGACGATGAG 180
Db 143 AAGGTCAACTTTCTCCATTTTATTTCAGGCAATTCCTTATGCTCCCATAGACGATGAG 202
Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTACCTTGTATATAAAGAAAGCG 240
Db 203 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTACCTTGTATATAAAGAAAGCG 262
Qy 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAATTAGA 300
Db 263 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAATTAGA 322
Qy 301 GAAAAATCTATTTTCAAGACCAAGAGAGACAAAGAGCTTACAGAGAGCAAGAGCTGCA 360
Db 323 GAAAAATCTATTTTCAAGACCAAGAGAGACAAAGAGCTTACAGAGAGCAAGAGCTGCA 382
Qy 361 GCAAGCGGGAGATCAAAAAATACGCACTAAGTGTCTATGATGAAGATTGA 410
Db 383 GCAAGCGGGAGATCAAAAAATACGCACTAAGTGTCTATGATGAAGATTGA 432

RESULT 7

BU567908
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU567908 823 bp mRNA linear EST 16-SEP-2002
AGENCOURT_10398893 NIH_MGC_82 Homo sapiens CDNA clone IMAGE:6614647
5', mRNA sequence.
BU567908
BU567908.1 GI:22918208
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria;
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2855 row: h column: 07
High quality sequence stop: 540.
Location/Qualifiers

FEATURES

1. 823
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6614647"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcttcggcc); Site 2: SfiI (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 32.1%; Score 405; DB 5; Length 823;
Best Local Similarity 100.0%; Pred. No. 9.3e-196; Mismatches 0; Indels 0; Gaps 0;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGCAGAGACTGCGGTCTTTCTGTCT 60
Db 52 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGCAGAGACTGCGGTCTTTCTGTCT 111
Qy 61 CTGCCCCCTCAAGGCGTGTGCTCAGACACGACGCTTCTGACGGAATACTATCTG 120
Db 112 CTGCCCCCTCAAGGCGTGTGCTCAGACACGACGCTTCTGACGGAATACTATCTG 171
Qy 121 AAGGTCAACTTTCTCCATTTTATTTCAGGCAATTCCTTATGCTCCCATAGACGATGAG 180
Db 172 AAGGTCAACTTTCTCCATTTTATTTCAGGCAATTCCTTATGCTCCCATAGACGATGAG 231
Qy 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTACCTTGTATATAAAGAAAGCG 240
Db 232 AGCAGCAAGCAAGATTGGGAATGACACCATTTGTCTTACCTTGTATATAAAGAAAGCG 291
Qy 241 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAATTAGA 300
Db 292 GCCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATCAAGAATTAGA 351
Qy 301 GAAAAATCTATTTTCAAGACCAAGAGAGACAAAGAGCTTACAGAGCAAGAGCTGCA 360
Db 352 GAAAAATCTATTTTCAAGACCAAGAGAGACAAAGAGCTTACAGAGCAAGAGCTGCA 411

```

QY 361 GCAAGCGGAAGATCAAAATACGCACTAAGTGTATGATGAAG 405
Db 412 GCAAGCGGAAGATCAAAATACGCACTAAGTGTATGATGAAG 456

RESULT 8
BG771796
LOCUS 792 bp mRNA linear EST 15-MAY-2001
DEFINITION 602720472F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837505 5',
mRNA sequence.
ACCESSION BG771796
VERSION BG771796.1 GI:14082449
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 792)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10769 row: p column: 18
High quality sequence stop: 671.
Location/Qualifiers
1..792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4837505"
/lab_host="DH10B"
/clone_lib="NIH_MGC 97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcagg); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 29.9%; Score 378; DB 4; Length 792;
Best Local Similarity 99.8%; Pred. No. 6.4e-182;
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCTCTTCAGGTACGATTACAGCTGGCAGCAGCAAGACTGCGGTCTTCTCTCT 60
Db 37 ATGCCTCTTCAGGTACGATTACAGCTGGCAGCAGCAAGACTGCGGTCTTCTCTCT 96

QY 61 CTGCCCTTCAAGCGGTGCGTTCAGACACGACGCTGTTCTGCAGGAAACTATCTG 120
Db 97 CTGCCCTTCAAGCGGTGCGTTCAGACACGACGCTGTTCTGCAGGAAACTATCTG 156

QY 121 AAGTGTCAACTTCTCCATTTTATTTTGGGCACTTCTTATGCTCCCATAGCAGTAG 180
Db 157 AAGTGTCAACTTCTCCATTTTATTTTGGGCACTTCTTATGCTCCCATAGCAGTAG 216

QY 181 AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCACTTGTATATAAAGAGCG 240
Db 217 AGCAGCAAGCAAGATTGGGAATGACACCATTTCTTCACTTGTATATAAAGAGCG 276

QY 241 GCCATGTGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGATTAGA 300
Db 277 GCCATGTGGAGACCCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAGATTAGA 336

QY 301 GAAAAATCTATTTTACAAGCACAAAGAGAGAGCAAAAAGAGCTACAGAAAGAGCTGCA 360
Db 337 GAAAAATCTATTTTACAAGCACAAAGAGAGAGCAAAAAGAGCTACAGAAAGAGCTGCA 396

QY 361 GCAAGCGGGAGATCAAAAATACGCACTAAGTGTATGATGAAGATTGAAGAAGAGAG 420
Db 397 GCAAGCGGGAGATCAAAAATACGCACTAAGTGTATGATGAAGATTGAAGAAGAGAG 456

QY 421 AGCAAAAAA 429
Db 457 AGCAAAAAA 465

RESULT 9
CD358543
LOCUS 933 bp mRNA linear EST 29-MAY-2003
DEFINITION AGENCOURT_14255831 NIH_MGC_180 Homo sapiens cDNA clone
IMAGE:30386203 5', mRNA sequence.
ACCESSION CD358543
VERSION CD358543.1 GI:31129978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 933)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM454 row: h column: 20
High quality sequence start: 303
High quality sequence stop: 666.
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30386203"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC 180"
/notes="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 29.6%; Score 374; DB 6; Length 933;
Best Local Similarity 99.8%; Pred. No. 7.3e-180;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 437 ATATGAAGAAATGACCGATAAAGCCCACTAAGCATTTGGAGCCCTGGAAGATATC 496
Db 313 ATATGAAGAAATGACCGATAAAGCCCACTAAGCATTTGGAGCCCTGGAAGATATC 372

```


/lab host="DH10B (Life Technologies)"
 /clone_lib="NCI_CGAP_FLO"
 /notes="Organ: Chondrosarcoma; Vector: pTT3-Pac
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP FLO is a cDNA library derived from
 a pool of mRNA obtained from 4 cell lines from grade III
 chondrosarcoma tissues. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pTT3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GAGTGGTG. The cell line
 was provided by Dr James Martin from University of Iowa.
 TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG_LIB=UI-H-FLO
 TAG_SEQ=GAGTGGTG"

ORIGIN

Query Match 27.1%; Score 342; DB 6; Length 772;
 Best Local Similarity 100.0%; Pred. No. 1.8e-163; Indels 0; Gaps 0;
 Matches 342; Conservative 0; Mismatches 0;
 Qy 922 GCAGCTATCATGATATAATTTAGCCATAAGACTAAATAATAGATGCCACTATTGTAT 981
 Db 720 GCAGCTATCATGATATAATTTAGCCATAAGACTAAATAATAGATGCCACTATTGTAT 661
 Qy 982 TTGAACCGGGCTGCTGCCACCTAAACTAAAACTTACCAAGGCTATTGAAGATTCT 1041
 Db 660 TTGAACCGGGCTGCTGCCACCTAAACTAAAACTTACCAAGGCTATTGAAGATTCT 601
 Qy 1042 TCTAAGCACTGGAAATTATTCATGCCACCTGTTACAGCAATGCTAATGCAAGAAATGAAG 1101
 Db 600 TCTAAGCACTGGAAATTATTCATGCCACCTGTTACAGCAATGCTAATGCAAGAAATGAAG 541
 Qy 1102 GCACATGTACGACGTGGAACAGCACTTCTGCACTAGAAATGTTAGTAGAGGCTACAG 1161
 Db 540 GCACATGTACGACGTGGAACAGCACTTCTGCACTAGAAATGTTAGTAGAGGCTACAG 481
 Qy 1162 GATTATGAGCGGCACATTAGATTGATCCATCCACAAATTTGACAAATTTGATGCTGAG 1221
 Db 480 GATTATGAGCGGCACATTAGATTGATCCATCCACAAATTTGACAAATTTGATGCTGAG 421
 Qy 1222 AAGATTGCGATGTAATTTCAAGGAACAGAACTAAATCTTAA 1263
 Db 420 AAGATTGCGATGTAATTTCAAGGAACAGAACTAAATCTTAA 379

RESULT 12

CE297042
 LOCUS
 DEFINITION CB297042 468 bp mRNA linear EST 28-FEB-2003
 12B22058_rev_1_F09_r_075.ab1 Chimpanzee brain library Koo's Pan
 troglodytes cDNA clone 12B22058_rev_1_F09_r_075.ab1 5', mRNA
 sequence.
 CB297042
 CB297042.1 GI:28622472
 VERSION
 EST.
 KEYWORDS
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (bases 1 to 468)
 Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B. and
 Paabo, S.
 Selection on human genes as revealed by comparisons to chimpanzee
 cDNA
 TITLE
 JOURNAL Genome Res. (2003) In press
 COMMENT Contact: Paabo S
 Evolutionary Genetics
 Max-Planck-Institute for evolutionary Anthropology

Deutscher Platz 6, 04103 Leipzig, Germany
 Tel: +49 (0)-341-3550 500
 Fax: +49 (0)-341-3550 555
 Email: paabo@eva.mpg.de
 Seq primer: M13 reverse.

FEATURES
 source

Location/Qualifiers
 1..468
 /organism="Pan troglodytes"
 /mol_type="mRNA"
 /db_xref="taxon:9598"
 /clone="12B22058_rev_1_F09_r_075.ab1"
 /sex="male"
 /tissue type="brain, presumably cortex"
 /dev stage="adult"
 /lab_host="Epichurian Coli (TM) XL-10-Gold"
 /clone_lib="Chimpanzee brain library Koo's"
 /notes="Vector: pUCHi; Site 1: SfiI-A; Site 2: SfiI-B; The
 library was prepared using the SMART cDNA library
 construction kit (Clontech), doing only primer extension,
 but not PCR amplification of the cDNA. The only deviation
 from the published protocol was that we cloned the cDNA
 into a plasmid vector."

ORIGIN

Query Match 26.9%; Score 340; DB 6; Length 468;
 Best Local Similarity 99.7%; Pred. No. 1.9e-162; Indels 0; Gaps 0;
 Matches 390; Conservative 0; Mismatches 1;
 Qy 49 GTCTTTCTGTCTCTGCCCTCAAGGCGTGTGGTCTAGAGACGCGGCTGTTCTGCACG 108
 Db 78 GTCTTTCTGTCTCTGCCCTCAAGGCGTGTGGTCTAGAGACGCGGCTGTTCTGCACG 137
 Qy 109 GAAACTATCTGAAGTCAACTTTCTCCATTTTATTATTTAGGCAATTTCTTATGCTCCC 168
 Db 138 GAAACTATCTGAAGTCAACTTTCTCCATTTTATTATTTAGGCAATTTCTTATGCTCCC 197
 Qy 169 ATAGACCATGAGAGCAGCAAGCAAGATTTGGGAATGACACCATTTGTTCAAGAGATGATG 228
 Db 198 ATAGACCATGAGAGCAGCAAGCAAGATTTGGGAATGACACCATTTGTTCAAGAGATGATG 257
 Qy 229 AAAAAAGAGCGGCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATG 288
 Db 258 AAAAAAGAGCGGCATGTGGGAGACCCCTTTCTGTGACGGGTGTGACAAAGAGATGATG 317
 Qy 289 CAAAGATTAGAGAAAAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTACAGAA 348
 Db 318 CAAAGATTAGAGAAAAATCTATTTTACAGCAACAGAGAGAGCAAAAGAGCTACAGAA 377
 Qy 349 GCAAAAGCTGCAGCAAGCGGAGAGATCAAAAATACCACTTAAGTGTATGATGAAGATT 408
 Db 378 GCAAAAGCTGCAGCAAGCGGAGAGATCAAAAATACCACTTAAGTGTATGATGAAGATT 437
 Qy 409 GAAGAAGAGAGAGAGAAAAAATAGAGAGATA 439
 Db 438 GAAGAAGAGAGAGAGAAAAAATAGAGAGATA 468

RESULT 13

AI024221
 LOCUS
 DEFINITION ov81e02.sl Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1643738
 3', mRNA sequence.
 AI024221
 AI024221.1 GI:3239265
 VERSION
 EST.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 417)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40ml3 fwd. ET from Amerham.

Location/Qualifiers
 1. .417
 /organism="Homo sapiens"
 /mol_type="mrna"
 /db_xref="taxon:9606"
 /clone="IMAGE:1643738"
 /sex="male"
 /lab_host="DH10B"
 /clone_lib="Soares testis_NHT"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer 15".
 TGTGACCAATCGAAGTGGGCGGCCGCCCAATTTTTTTTTTTT 3'.
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 26.8%; Score 338; DB 1; Length 417;
 Best Local Similarity 99.7%; Pred. No. 2e-161;
 Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 149 AGGCATTTCTTTATGCTCCATAGACGATGAGCAGCAAGCAAAAGATTGGGAATGACA 208
 DB 1 AGGCATTTCTTTATGCTCCATAGACGATGAGCAGCAAGCAAAAGATTGGGAATGACA 60
 QY 209 CAAATGCTCTTACCTTGAT 268
 DB 61 CCAATGCTCTTACCTTGAT 120
 QY 269 GTGTTGACAAAGAGATGATGCAAGCAATTTAGAGAAAATCTATTTTACAGCAACAGAGA 328
 DB 121 GTGTTGACAAAGAGATGATGCAAGCAATTTAGAGAAAATCTATTTTACAGCAACAGAGA 180
 QY 329 GAGCAAAAGAACTTACAGAAAGCAAAAGCTGCAGCAAAAGCGGGAAGATCAAAAATATCGCAC 388
 DB 181 GAGCAAAAGAACTTACAGAAAGCAAAAGCTGCAGCAAAAGCGGGAAGATCAAAAATATCGCAC 240
 QY 389 TAAGTGTCTATGATGAAGATTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
 DB 241 TAAGTGTCTATGATGAAGATTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 449 ATGAACGGATAAAGCCACTAAAGCATTTGGAGCTGGAGAGATATCAAGAGAAAGCTG 508
 DB 301 ATGAACGGATAAAGCCACTAAAGCATTTGGAGCTGGAGAGATATCAAGAGAAAGCTG 360
 QY 509 AGGAGCAAAAAAAATTTACAGAGAGAGAG 537
 DB 361 AGGAGCAAAAAAAATTTACAGAGAGAGAG 389

RESULT 14
 AI783611/c
 LOCUS
 DEFINITION
 similar to contains Alu repetitive element1, mRNA sequence.
 ACCESSION
 AI783611.1 GI:5325420

EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 325)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 437 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 325.
 Location/Qualifiers
 1. .325
 /organism="Homo sapiens"
 /mol_type="mrna"
 /db_xref="taxon:9606"
 /clone="IMAGE:2296704"
 /lab_host="DH10B"
 /clone_lib="NCI-CGAP Kid11"
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP Kid11 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNA from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 25.7%; Score 325; DB 1; Length 325;
 Best Local Similarity 100.0%; Pred. No. 8.9e-155;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 TTCAGAGAGAGAGAAATTTATGTCAAAAGAAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAG 583
 DB 325 TTCAGAGAGAGAGAAATTTATGTCAAAAGAAAAGCAAAATTAAGAGAGAGAGAGAGAGAGAG 266
 QY 584 TAAATATATAGAGCTTACTAGAAAATTTGGCATCTAGAAAATCTTGTCTCCAAAAGGGAGAA 643
 DB 265 TAAATATATAGAGCTTACTAGAAAATTTGGCATCTAGAAAATCTTGTCTCCAAAAGGGAGAA 206
 QY 644 ATTCAAAAAATATATTTACTCAGAGAGTTTAAAGAGAGAGAGAGAGAGATTTCTCGCTCTG 703
 DB 205 ATTCAAAAAATATATTTACTCAGAGAGTTTAAAGAGAGAGAGAGATTTCTCGCTCTG 146
 QY 704 TTGGCAGTATTAATCACTTTTACCTTCAGATTTCCCAACAGCTCTTCGTGAATCAC 763
 DB 145 TTGGCAGTATTAATCACTTTTACCTTCAGATTTCCCAACAGCTCTTCGTGAATCAC 86
 QY 764 AGTAGCAG 823
 DB 85 AGTAGCAG 26
 QY 824 CTGACATAGCTGAACCTTTGCGATTT 848
 DB 25 CTGACATAGCTGAACCTTTGCGATTT 1

RESULT 15

BUS52580
LOCUS
DEFINITION ACENOCURT_10501009 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6619012
5', mRNA sequence.
ACCESSION BUS52580
VERSION BUS52580.1 GI:24037543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 651)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2866 row: n column: 04
High quality sequence stop: 364.
FEATURES
source
1..651
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6619012"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site:1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCATTATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGCCGCGGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
ORIGIN
Query Match 25.6%; Score 323; DB 5; Length 651;
Best Local Similarity 99.7%; Pred. No. 1e-153;
Matches 373; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 875 GGTGAAGGATAAAGGAACAATAATTGTTGCAACGGAAACTATTGGCAGCTATCAATG 934
Db 3 GGTGAAGGATAAAGGAACAATAATTGTTGCAACGGAAACTATTGGCAGCTATCAATG 62
Qy 935 CATATAATTTAGCCATAGACTAATAATAAGATGCCACTATTGTATTGAACCGGCTG 994
Db 63 CATATAATTTAGCCATAGACTAATAATAAGATGCCACTATTGTATTGAACCGGCTG 122
Qy 995 CTTCGCCACCTAAAACTAAAAAATTTACACAGGCTATTGAAGATTCCTTCTAAGGCACCTGG 1054
Db 123 CTGCACCTAAAACTAAAAAATTTACACAGGCTATTGAAGATTCCTTCTAAGGCACCTGG 182
Qy 1055 AATTATTGATGCCACCTGTTACAGACATGCTAATGCAAGATGAAGGCACATGTACGAC 1114
Db 183 AATTATTGATGCCACCTGTTACAGACATGCTAATGCAAGATGAAGGCACATGTACGAC 242
Qy 1115 GTGGAACAGCAATTCGTCAACTAGAAATTTGATGTAGAGGCCTCAGGATTATCAAGCGG 1174
Db 243 GTGGAACAGCAATTCGTCAACTAGAAATTTGATGTAGAGGCCTCAGGATTATCAAGCGG 302
Qy 1175 CACTTAAGATTGATCCATCCAAACAAAATTTGACAAAATTTGATGCTGAGAAGATTCGGAATG 1234
Db 303 CACTTAAGATTGATCCATCCAAACAAAATTTGACAAAATTTGATGCTGAGAAGATTCGGAATG 362

Qy 1235 TAATTCAAGGAACA 1248
|||||
Db 363 TAATTCAAGGAACA 376
|||||

Search completed: July 8, 2005, 12:03:37
Job time : 9626 secs